PATENT ABSTRACTS OF JAPAN

(11)Publication number : 10-124089

(43)Date of publication of application : 15.05.1998

(51)Int.Cl. G10L 3/02
G10L 9/08
H04M 1/00

(21)Application number : 08-282235

(22)Date of filing : 24.10.1996

(71)Applicant : SONY CORP

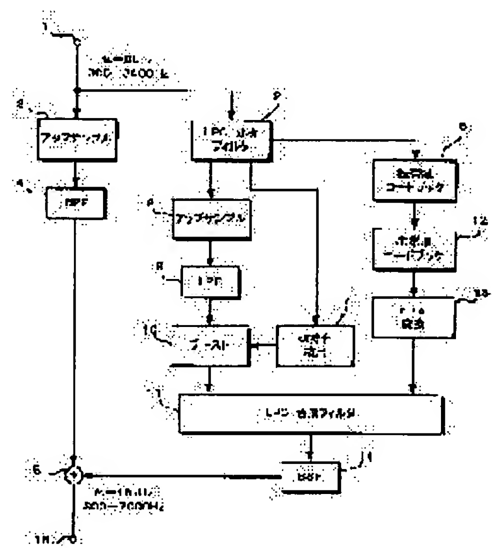
(72)Inventor : OMORI SHIRO
NISHIGUCHI MASAYUKI

(54) PROCESSOR AND METHOD FOR SPEECH SIGNAL PROCESSING AND DEVICE AND METHOD FOR EXPANDING VOICE BANDWIDTH

(57)Abstract:

PROBLEM TO BE SOLVED: To clearly reproduce a fricative and an affricate when voice bandwidth is expanded by LPC synthesis.

SOLUTION: An input voice signal is analyzed by an LPC analytic filter 2 to retrieve the most matching autocorrelation as a parameter from a narrow- band code book 6, the parameter corresponding to it is outputted from a broadband code book 12, and an LPC synthesizing filter 11 performs synthesis. Consequently, the bandwidth of a voice is expanded. An affricate detecting circuit 7 detects a fricative and an affricate by using the autocorrelation value of the input voice signal and the value of frame power. Once the fricative or affricate is detected, the whole or part of the band of an exciting source is boosted. Consequently, a power deficiency in case of the fricative or affricate is improved to clearly reproduce the fricative and affricate.



LEGAL STATUS

[Date of request for examination]

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]

* NOTICES *

CLAIMS

[Claim(s)]
[Claim 1] The sound signal processor carry out having carried out as [have / an affricate detection means detect the fricative and affricate of the above-mentioned input sound signal, and a boost means give a boost to the source of excitation when the above-mentioned fricative and the affricate are detected] as the description in the sound signal processor which compounded the sound signal after analyzing an input sound signal and performing signal processing to the sound signal by which analysis was carried out [above-mentioned].
[Claim 2] The above-mentioned affricate detection means is a sound signal processor according to claim 1 which is what detects a fricative and the affricate using the value of the autocorrelation of the above-mentioned input sound signal, and the value of frame power at least.
[Claim 3] The above-mentioned boost means is a sound signal processor according to claim 1 or 2 which is what changes a boost value to ****.

[Claim 4] The sound signal art characterized by carrying out as [boost / all the bands of the source of excitation, or some bands] when the fricative and affricate of the above-mentioned input sound signal are detected and the above-mentioned fricative and the affricate are detected in the sound signal art which compounded the sound signal after analyzing the input sound signal and performing signal processing to the sound signal by which analysis was carried out [above-mentioned].
[Claim 5] Detection of the above-mentioned affricate is the sound signal art according to claim 4 which was made to perform using the value of the autocorrelation of the above-mentioned input sound signal, and the value of frame power at least.
[Claim 6] The sound signal art according to claim 4 or 5 it was made to change the above-mentioned boost value to ****.
[Claim 7] An analysis means to ask for a parameter from an input narrow-band sound signal, and the source means forming of excitation which asks for the source of excitation from the LPC remainder of the above-mentioned input narrow-band sound signal, The narrow-band code book with which the parameter of the narrow-band sound signal beforehand acquired from the pattern of two or more sound signals was stored, The broadband code book with which the parameter of the wideband voice signal beforehand acquired from the pattern of two or more sound signals was stored corresponding to the above-mentioned narrow-band code book, An affricate detection means to detect a fricative and the affricate, and a boost means to give a boost to the above-mentioned source of excitation when the above-mentioned fricative and the affricate are detected, A matching means to compare the parameter of the sound signal of the above-mentioned input narrow-band with the parameter of the input narrow-band sound signal stored in the above-mentioned narrow-band code book, and to search the optimal parameter, Based on the retrieval result in the above-mentioned matching means, the parameter which corresponds out of the parameter of the wideband voice signal stored in the above-mentioned broadband code book is read. Speech bandwidth growth equipment characterized by having the parameter by which reading appearance was carried out [above-mentioned], and a synthetic means to compound an output wideband voice signal based on the above-mentioned source of excitation.
[Claim 8] The above-mentioned affricate detection means is speech bandwidth growth equipment according to claim 7 which is what detects a fricative and the affricate using the value of the autocorrelation of the above-mentioned input sound signal, and the value of frame power at least.
[Claim 9] The above-mentioned boost means is speech bandwidth growth equipment according to claim 7 or 8 which is what changes a boost value to ****.
[Claim 10] The narrow-band code book with which the parameter of the narrow-band sound signal beforehand acquired from the pattern of two or more sound signals was stored, The broadband code book with which the parameter of the wideband voice signal beforehand acquired from the pattern of two or more sound signals was stored corresponding to the above-mentioned narrow-band code book is prepared. In quest of a parameter, analyze from an input narrow-band sound signal, and it asks for the source of excitation from the LPC remainder of the above-mentioned input narrow-band sound signal. When a fricative and the affricate are detected and the above-mentioned fricative and the affricate are detected, a boost is given to the above-mentioned source of excitation. The parameter of the sound signal of the above-mentioned input narrow-band, The parameter of the input narrow-band sound signal stored in the above-mentioned narrow-band code book is compared. Search the optimal parameter and it is based on the retrieval result in the above-mentioned matching. The speech bandwidth escape approach characterized by reading the parameter which corresponds out of the parameter of the wideband voice signal stored in the above-mentioned broadband code book, and compounding an output wideband voice signal based on the parameter and the above-mentioned source of excitation by which reading appearance was carried out [above-mentioned].
[Claim 11] The above-mentioned affricate and a fricative are the speech bandwidth escape approach according to claim 10 which is what is detected using the value of the autocorrelation of the above-mentioned input sound signal, and the value of frame power.
[Claim 12] The speech bandwidth escape approach according to claim 10 or 11 of having made it change the above-mentioned boost value to ****.

[Translation done.]

* NOTICES *

DETAILED DESCRIPTION

[Detailed Description of the Invention]
[0001]
[Field of the Invention] This invention relates to bandwidth growth equipment and an approach by minding transmission lines, such as the telephone line, from the sound signal with which the frequency band is restricted to the narrow-band at the sound signal processor for generating the sound signal of a broadband and an approach, and a list.
[0002]
[Description of the Prior Art] The band of the telephone line is as narrow as 300-3400kHz, and the frequency band of the sound signal sent through the telephone line is restricted. For this reason, the tone quality of the conventional analog telephone line can seldom be said to be fitness. Moreover, there is dissatisfaction also about the tone quality of a digital cellular phone.
[0003] Then, speech bandwidth is extended by the receiver side and the system which aimed at the improvement of tone quality is proposed variously. The narrow-band code book with which the parameter of the narrow-band sound signal beforehand acquired from the pattern of two or more sound signals in this was stored as a code vector, The broadband code book with which the parameter of the wideband voice signal acquired from the pattern of the same sound signal as this was beforehand stored as a code vector is prepared. By a narrow-band code book's analyzing an input signal, and synthesizing voice using a broadband code book based on this analysis result, speech bandwidth is extended and the system which improved tone quality is proposed.
[0004] That is, a frequency band is restricted [as shown in drawing 6 ,] when transmitting a sound signal through a transmission line like the telephone line, and the sound signal from transmission side 101 minds a transmission line 102. For example, even if there is about 7000Hz of frequency bands of the sound signal from transmission side 101 from 300Hz, the frequency band of the sound signal sent to receiver side 103 is restricted to about 3400Hz from 300Hz by minding a transmission line 102.
[0005] Then, as shown in drawing 7 , the narrow-band code book 105 with which the parameter of the narrow-band sound signal beforehand acquired from the pattern of two or more sound signals was stored as a code vector, and the broadband code book 106 with which the parameter of the wideband voice signal with which it was obtained from the pattern of the same sound signal corresponding to the narrow-band code book 105 was beforehand stored as a code vector are prepared.
[0006] In addition, code books 105 and 106 divide the sound signal of the same broadband into the frame of predetermined die length, for example, form the pattern of two or more sound signals, and are created by analyzing spectrum envelopment for every frame. Namely, the sound signal of a broadband is used for code book creation time, and the sound signal of this broadband is divided into it for every predetermined frame. The spectrum envelopment information when analyzing the sound signal of this broadband with a broadband is stored in the broadband code book 106 as a code vector. The spectrum envelopment information when band-limiting the sound signal of a broadband to 300-3400Hz, and analyzing it is stored in the narrow-band code book 105 as a code vector.
[0007] As spectrum envelopment information stored in the narrow-band code book 105 and the broadband code book 106, LPC KEPUTORAMU is used conventionally. LPC KEPUTORAMU is KEPUTORAMU by linear predictor coefficients, is shown by the following formulas, and is made and called for.
[Equation 1]

$$\begin{cases} c_1 = -\alpha_1 \\ c_n = -\alpha_n - \sum_{m=1}^{n-1} \left(1 - \frac{m}{n}\right) \alpha_m c_{n-m} \qquad (1 < n \leqq p) \\ c_n = -\sum_{m=1}^{p} \left(1 - \frac{m}{n}\right) \alpha_m c_{n-m} \end{cases}$$

α ：線形予測係数
p ：線形予測次数

[0008] In drawing 7 , the sound signal of the narrow-band sent to receiver side 103 from transmission side 101 through a transmission line 102 is first sent to the analysis circuit 104. An input sound signal is divided for every predetermined frame, and spectrum envelopment is called for in the analysis circuit 104. The output of the analysis circuit 104 is sent to the narrow-band code book 105. With the narrow-band code book 105, the spectrum envelopment analyzed in the analysis circuit 104 is compared with the spectrum envelopment information stored in the narrow-band code book 105, and matching processing is performed. And the output of the narrow-band code book 105 is sent to the broadband code book 106, and the spectrum envelopment information which matches most in the narrow-band code book 105, and the spectrum envelopment information on a broadband that it corresponds are read from the broadband code book 106.
[0009] This broadband spectrum envelopment information is sent to the synthetic circuit 107. A sound signal is compounded in the synthetic circuit 107 using the spectrum envelopment information on the broadband read from the broadband code book 106. Since this compounded sound signal is compounded using the broadband code book 106, it turns into a sound signal of a broadband.

[0010]
[Problem(s) to be Solved by the Invention] As mentioned above, in the conventional speech bandwidth escape system, LPC KEPUTORAMU is used as a code book vector. Moreover, the noise and the pulse train are used as a source of excitation at the time of compounding a sound signal. however, the case where the distortion on audibility and a quantization error use the linear scale of a match comparatively in LPC KEPUTORAMU -- a logarithm -- since a scale is used, the small part of energy is thought as important and the error in the large part of energy becomes large. In order to use for such a speech bandwidth escape system, on audibility, it is desirable to suppress distortion by the vowel part as much as possible. Therefore, LPC KEPUTORAMU cannot necessarily be said to be the optimal thing. Moreover, although the thing possible nearest to the LPC remainder of a broadband must be good about the source of excitation, the conventional method using a noise and a pulse train is far from this.
[0011] Then, it is possible [it] to compound a wideband voice signal by LPC composition, using what carried out the rise sample of the LPC remainder as a source of excitation, using an autocorrelation as a code book vector. an autocorrelation -- a logarithm -- since it is not a scale, it is thought that distortion by the vowel part is improved. However, if the sound signal of a broadband is formed by LPC composition, using what carried out the rise sample of the LPC remainder as a source of excitation, using an autocorrelation as a code book vector, especially, a fricative and the affricate will run short and the problem of becoming an unclear sound will arise. It is thought that this originates in the lack of power of the source of excitation as a seed although it is raised to a cause that prediction of spectrum envelopment is not enough, either.
[0012] That is, in the case of a fricative and the affricate, prediction by LPC composition is performed comparatively well, and the power of the remainder becomes small. However, wideband voice of prediction is inadequate and the power of the remainder does not become small. For this reason, in case the band of a fricative and the affricate is extended, remainder power must also be large equally with it. However, since the remainder is predicted and created from the narrow-band remainder, power is not large enough. For this reason, in the case of a fricative and the affricate, the power of the source of excitation runs short.
[0013] Therefore, in case the purpose of this invention extends speech bandwidth, it is to provide with speech bandwidth growth equipment and an approach the sound signal processor which enabled it to reproduce a fricative and the affricate clearly and an approach, and a list.
[0014]
[Means for Solving the Problem] After this invention analyzes an input sound signal and performs signal processing to the analyzed sound signal, it is a sound signal processor characterized by to carry out as [have / an affricate detection means detect the fricative and affricate of an input sound signal in the sound signal processor which compounded the sound signal, and a boost means give a boost to the source of excitation when a fricative and the affricate are detected].
[0015] Moreover, an analysis means by which this invention asks for the parameter of an autocorrelation from an input narrow-band sound signal, The source means forming of excitation which asks for the source of excitation from the LPC remainder of an input narrow-band sound signal, The narrow-band code book with which the parameter of the autocorrelation of the narrow-band sound signal beforehand acquired from the pattern of two or more sound signals was stored, The broadband code book with which the parameter of the autocorrelation of the wideband voice signal beforehand acquired from the pattern of two or more sound signals was stored corresponding to the narrow-band code book, An affricate detection means to detect a fricative and the affricate, and a boost means to give a boost to the source of excitation when a fricative and the affricate are detected, A matching means to compare the parameter of the autocorrelation of the sound signal of an input narrow-band with the parameter of the autocorrelation of the input narrow-band sound signal stored in the narrow-band code book, and to search the optimal parameter, Based on the retrieval result in a matching means, the parameter which corresponds out of the parameter of the autocorrelation of the wideband voice signal stored in the broadband code book is read. It is speech bandwidth growth equipment characterized by having a synthetic means to compound an output wideband voice signal based on this read parameter and source of excitation.
[0016] In this invention, an affricate detection means detects a fricative and the affricate using the value of the autocorrelation of an input sound signal, and the value of frame power.
[0017] Thus, if it is made ** which gives a boost to the source of excitation when [which detects a fricative and the affricate] affricate detection is carried out and a fricative and the affricate are detected, the lack of power in the case of a fricative or the affricate is improved, and a fricative and the affricate can be reproduced clearly.
[0018]
[Embodiment of the Invention] Hereafter, the gestalt of implementation of this invention is explained with reference to a drawing. Drawing 1 shows an example of the speech bandwidth escape system by which this invention was applied. In drawing 1 , the narrow-band sound signal whose sampling frequency a frequency band is 8kHz in 300Hz - 3400Hz is supplied to an input terminal 1. This narrow-band sound signal is supplied to the rise sample circuit 3 while it is supplied to the LPC (Linear Predictive Coding) analysis filter 2.
[0019] The rise sample circuit 3 is for carrying out the rise sample of the sampling frequency to 16kHz from 8kHz. The output of the rise sample circuit 3 is supplied to an adder circuit 5 through the band pass filter 4 of a 300Hz - 3400Hz passband. The path leading to this rise sample circuit 3, a band pass filter 4, and an adder circuit 5 is a path for adding the signal of the component of the original frequency band to the sound signal of a high region from which it synthesized voice, as explained later.
[0020] The LPC analysis filter 2 frame-izes the narrow-band sound signal from an input terminal 1, and performs 10th LPC analysis. The 10th autocorrelation is obtained in process of LPC analysis. This autocorrelation is sent to the affricate detector 7 while it is sent to the narrow-band code book 6. Moreover, the LPC remainder searched for with the LPC analysis filter 2 is sent to the rise sample circuit 8.
[0021] The rise sample of the LPC remainder of the voice of a narrow-band is carried out by the rise sample circuit 8. The output of the rise sample circuit 8 is sent to the LPC composition filter 11 through a low pass filter 9 and boost circuit 10 **. The signal which carried out the rise sample of this LPC remainder, and oppressed the high region is used as a source of excitation at the time of compounding a sound signal so that it may explain later. The boost circuit 10 is for boosting the source of excitation, when the affricate and a fricative are detected, and the amount of boosts of the boost circuit 10 is controlled by the output of the affricate detector 7.
[0022] The 10th autocorrelation information on the narrow-band sound signal beforehand acquired from the pattern of two or more sound signals is stored in the narrow-band code book 6 as a code vector. With the narrow-band code book 6, the autocorrelation obtained from the LPC analysis filter 2 is compared with the autocorrelation information stored in the narrow-band code book 6, and matching processing is performed. And the index of the autocorrelation information which matches most is sent to the broadband code book 12.
[0023] Corresponding to the narrow-band code book 6, the 20th autocorrelation information on the wideband voice signal acquired from the sound signal of the same pattern as the time of creating the narrow-band code book 6 is stored in the broadband code book 12 as a code vector. If the autocorrelation information which matches most with the narrow-band code book 6 is judged, this index will be sent to the broadband code book 12, and the autocorrelation information on the broadband corresponding to the autocorrelation information on the narrow-band judged to match most with the broadband code book 12 will be read.

[0024] An autocorrelation is the parameter of a time domain, is the following, and is made and called for.
[Equation. 2]

$$r_{\tau} = \sum_{t=0}^{N-1-\tau} x_t x_{t-\tau} \qquad (\tau \geqq \phi)$$

$$\{x_t\} = \{x_0, x_1, \cdots x_{N-1}\}$$

N ：音声サンプル数

[0025] Using the wideband voice signal whose sampling frequency is 16kHz and which is 0-8000kHz, the broadband code book 12 is the following, and is made and created. That is, this wideband voice signal is divided into 20m frame in every second of advance for die-length 32 m seconds, and the creation time of the broadband code book 12 is asked for the 20th autocorrelation with each frame. A 8-bit code book is created by the GLA (General Lloyd Algorithm) algorithm using this. Let this be the broadband code book 4. Here, the frame number encoded by the i-th code vector of a broadband code book is Ai(ed).
[0026] The narrow-band code book 6 is the same sound signal as having created the broadband code book 12, and a sampling frequency is created using what was restricted to 300Hz - 3400Hz in a frequency band by 8kHz. It is divided into a frame at the time of day as the time of creating the broadband code book 12 when the sound signal restricted to this narrow-band is the same, and the 10th autocorrelation is called for with each frame. And the center of gravity of the narrow-band autocorrelation of the frame belonging to a frame number Ai is searched for, and it is made to make it correspond to the broadband autocorrelation of the broadband code book of a frame number Ai by making the vector into the i-th code vector of the code book of a narrow-band.
[0027] In drawing 1 , the autocorrelation information on the broadband read from the broadband code book 12 is sent to the autocorrelation-linear-predictor-coefficients conversion circuit 13. Conversion to linear predictor coefficients from an autocorrelation is performed by the autocorrelation-linear-predictor-coefficients conversion circuit 13. These linear predictor coefficients are sent to the LPC composition filter 11.
[0028] The rise sample of the LPC remainder from the LPC analysis filter 2 is carried out in the rise sample circuit 8, distortion is generated in the LPC composition filter 11 by return, and the signal which oppressed the high region side through the low pass filter 9 is supplied to it. With the LPC composition filter 11, the rise sample of this LPC remainder is carried out, and LPC composition is performed by the linear predictor coefficients from the autocorrelation-linear-predictor-coefficients conversion circuit section 13, using what oppressed the distorted high region side by return as a source of excitation. Thereby, the sound signal of a 300Hz - 7000Hz broadband is compounded.
[0029] The sound signal compounded with the LPC composition filter 11 is supplied to a band stop filter 14. A band stop filter 14 removes the signal component of the frequency band of an input narrow-band sound signal. By the band stop filter 14, the 300Hz - 3400Hz signal component contained in the sound signal of the original narrow-band is removed out of the sound signal of a broadband with a frequency of 300Hz - 7000Hz compounded with the LPC composition filter 11. The output of this band stop filter 14 is supplied to an adder circuit 5.
[0030] In an adder circuit 5, the component of the sound signal of the narrow-band of origin with a frequency [through the rise sample circuit 3 and a band pass filter 4] of 300Hz - 3400Hz and the component of the sound signal with a frequency [through a band stop filter 14] of 3400Hz - 7000Hz from which it synthesized voice are added. Thereby, the digital sound signal whose sampling frequency a frequency band is 16kHz in 300-7000Hz is acquired. This digital sound signal is outputted from an output terminal 15.
[0031] Thus, in the speech bandwidth growth equipment to which this invention was applied, an input narrow-band sound signal is analyzed using the narrow-band code book 6, and the sound signal of a broadband is compounded using the broadband code book 12. And an autocorrelation is used as information on a code book. rather than LPC KEPUTORAMU uses as spectrum envelopment information and ******** generally uses LPC KEPUTORAMU conventionally as a result of an experiment -- a logarithm -- it is because it turned out that it is more desirable on audibility to use the autocorrelation which is not a scale. this -- LPC KEPUTORAMU -- a logarithm -- since the scale is used -- the small consonant of power -- in a part, it is thought that it is because the error in the large vowel part of power becomes large relatively although an error becomes small.
[0032] And in the speech bandwidth escape system to which this invention was applied, as a source of excitation, the rise sample of the LPC remainder is carried out, distortion is generated by return, and what oppressed the distorted high region side by return is used. If it does in this way, since original audio power and harmonic structure are saved, engine performance sufficient as a source of excitation is obtained.
[0033] Thus, the sound signal of a 300Hz - 7000Hz good broadband is acquired from the LPC composition filter 11 by carrying out the rise sample of the LPC remainder, using as a source of thing excitation which oppressed the distorted high region side by return, using an autocorrelation as information on code books 6 and 12, and compounding a sound signal.
[0034] Thus, the sound signal of the broadband obtained from the LPC composition filter 11 also includes the signal of the frequency component of the original band, and if the output signal of the LPC composition filter 11 is used as it is in order that distortion may attain to the frequency component of the original band by these processings, the distorted effect of the frequency component of the original band will arise.
[0035] Then, by the band stop filter 14, from the output of the LPC composition filter 11, the frequency component of the band of the origin of 300Hz - 3400Hz is removed, and it is carrying out as [add / the component of the sound signal of the origin of 300Hz - 3400Hz taken out through the band pass filter 4, and the component of the 3400Hz - 7000Hz sound signal compounded with the LPC composition filter 11].
[0036] In addition, in distance count of code book creation time, it may be made to perform weighting processing so that the weight of high order data may become small. That is, in the narrow-band code book 6, weight from the 1st order to the 3rd order is set to "1", weight is set to "0" in the degree beyond it, weight from the 1st order to the 6th order is set to "1" in the broadband code book 12, and weight is set to "0" in the degree beyond it. If it does in this way, it not only can perform saving of memory space, but reappearance of a rough spectral envelope will be thought as important as a property of an autocorrelation parameter, and more quality voice will be obtained.
[0037] By the way, if the sound signal of a broadband is formed by LPC composition by making into the source of excitation what carried out the rise sample of the LPC remainder, and oppressed the high region in this way, using an autocorrelation as a code vector, especially, a fricative and the affricate will run short and it will become an unclear sound. It is thought that this originates mainly in the lack of power of the source of excitation although it is raised to a cause that prediction of spectrum envelopment is not enough, either.

[0038] So, in the system to which this invention was applied, the affricate detector 7 which detects a fricative and the affricate, and the boost circuit 10 which boosts all the bands of the source of excitation or some bands when a fricative and the affricate are detected are formed. The 10th autocorrelation called for with the LPC analysis filter 2 is supplied to the affricate detector 7. It is detected whether in the affricate detector 7, a fricative and the affricate were inputted among this 10th autocorrelation using zero-order frame power, the primary autocorrelation, and the secondary autocorrelation. When a fricative and the affricate are detected in the affricate detector 7, all the bands of the source of excitation or some bands are boosted by the boost circuit 10.
[0039] That is, the case of a vowel, and in the case of affricate [a fricative or], as a result of analyzing the autocorrelation of an input sound signal, it turned out that the following differences are in the physical relationship of a zero-order autocorrelation, i.e., frame power, the primary autocorrelation, and the secondary autocorrelation. That is, when R0 and the primary autocorrelation are made into R1 and the secondary autocorrelation R2, as zero-order frame power is shown in drawing 2 , when an input sound signal is a vowel, the zero-order frame power R0, the primary autocorrelation R1, and the secondary autocorrelation R2 are located in a line on an abbreviation straight line. On the other hand, as shown in drawing 3 , in the case of a fricative or the affricate, the physical relationship of the zero-order frame power R0, the primary autocorrelation R1, and the secondary autocorrelation R2 turns into relation which is located in a line convex. If it judges from this whether the physical relationship of R1 and the secondary autocorrelation R2 is located [power / zero-order / frame] in a line convex in R0 and the primary autocorrelation, detection of a fricative or the affricate can be performed.
[0040] When satisfied with the system to which this invention was applied using this of the following conditions, it is judged that they are a fricative and the affricate.
[0041] Conditions (1)
When R0 is more than constant value, R1 is more than constant value and R1/R2 are below constant value, it is judged that they are a fricative and the affricate.
[0042] Conditions (2)
In R's0 being below constant value more than constant value, and R's1 being below constant value and being 1-R1>R1-R2, it judges that they are a fricative and an explosive sound.
[0043] Conditions (3)
In R's0 being below constant value more than constant value, and (R1-dc) /'s (R0-dc's) being below constant value and being 1-R1>R1-R2, it judges that they are a fricative and an explosive sound. In addition, dc is a fixed value for every frame band.
[0044] When it is judged according to conditions (1) or conditions (2) that they are a fricative and the affricate, 10dB of sources of excitation is boosted, for example. Moreover, when it is judged according to conditions (3) that they are a fricative and the affricate, 5dB of sources of excitation is boosted, for example.
[0045] Moreover, if the source of excitation is boosted in an instant when the above conditions are fulfilled, a sound will change suddenly and sense of incongruity will be given. Then, he is made to carry out smoothing of the boost of the source of excitation for every frame, and is trying for change of a boost of the source of excitation not to be noticeable as the source of excitation does not change rapidly.
[0046] It is clear by experiment that the speech bandwidth escape of a good property is performed by the speech bandwidth escape system by which this invention was applied. That is, drawing 4 shows the experimental result when performing the bandwidth escape of a sound signal using the speech bandwidth escape system by which this invention was applied. Drawing 4 A is the spectrum Fig. of the sound signal of the broadband used as the source. The sound signal used as this source shall be band-limited as shown in drawing 4 B, and the speech bandwidth escape system by which this invention was applied shall perform a bandwidth escape. Drawing 4 C is the sound signal acquired by performing the bandwidth escape of this signal. If drawing 4 A is compared with drawing 4 C, the speech bandwidth escape system by which this invention was applied shows that the bandwidth escape of a sound signal was able to be performed in a remarkable precision.
[0047] In addition, this invention can be used for the tone-quality improvement of the telephone line of an analog, and a tone-quality improvement of a digital cellular phone. Especially, in the digital cellular phone, VSELP and PSI-CELP are used as a modulation technique. In VSELP or PSI-CELP, since linear predictor coefficients and the source of excitation are used, such information can be used in the case of the LPC analysis and LPC composition in a speech bandwidth escape system.
[0048] That is, drawing 5 shows the example of application in a digital cellular phone. As shown in drawing 5 , in a digital cellular phone, a parameter equivalent to the source of excitation, linear-predictor-coefficients alpha1 -alpha10, or this is sent. This source of excitation is supplied to an input terminal 21, and linear predictor coefficients are supplied to an input terminal 22. The source of excitation from an input terminal 21 is sent to the rise sample circuit 24 while it is sent to the LPC composition filter 23. The auto correlation coefficient from an input terminal 22 is sent to the LPC composition filter 23.
[0049] With the LPC composition filter 23, based on the source of excitation from an input terminal 21, the linear predictor coefficients from an input terminal 22 are used, and a sound signal is compounded. The sound signal compounded with the LPC composition filter 23 is supplied to the rise sample circuit 25.
[0050] The rise sample circuit 25 is for carrying out the rise sample of the sampling frequency. The output of the rise sample circuit 25 is supplied to an adder circuit 27 through a band pass filter 26. The path leading to this rise sample circuit 25, a band pass filter 26, and an adder circuit 27 is a path for adding to the sound signal which had the signal of the component of the original frequency band compounded.
[0051] Moreover, linear predictor coefficients are sent to the linear-predictor-coefficients-autocorrelation conversion circuit 28 from the LPC composition filter 23. The linear-predictor-coefficients-autocorrelation conversion circuit 28 changes linear predictor coefficients into an autocorrelation. This autocorrelation is sent to the affricate detector 30 while it is sent to the narrow-band code book 29.
[0052] Moreover, the source of excitation from an input terminal 21 is sent to the rise sample circuit 24. The output of the rise sample circuit 24 is sent to the LPC composition filter 33 through a low pass filter 31 and the boost circuit 32. The boost circuit 32 is for boosting the source of excitation, when the affricate and a fricative are detected, and the amount of boosts of the boost circuit 32 is controlled by the output of the affricate detector 30.
[0053] The autocorrelation information on the narrow-band sound signal beforehand acquired from the pattern of two or more sound signals is stored in the narrow-band code book 29 as a code vector. With the narrow-band code book 29, the autocorrelation from the linear-predictor-coefficients-autocorrelation conversion circuit 28 is compared with the autocorrelation information stored in the narrow-band code book 29, and matching processing is performed. And the index of the autocorrelation information which matches most is sent to the broadband code book 34.

[0054] Corresponding to the narrow-band code book 29, the autocorrelation information on the wideband voice signal acquired from the sound signal of the same pattern as the time of creating the narrow-band code book 29 is stored in the broadband code book 34 as a code vector. If the autocorrelation information which matches most with the narrow-band code book 29 is judged, this index will be sent to the broadband code

book 34, and the autocorrelation information on the broadband corresponding to the autocorrelation information on the narrow-band judged to match most with the broadband code book 34 will be read.

[0055] The autocorrelation information on the broadband read from the broadband code book 34 is sent to the autocorrelation-linear-predictor-coefficients conversion circuit 35. Conversion to linear predictor coefficients from an autocorrelation is performed by the autocorrelation-linear-predictor-coefficients conversion circuit 35. These linear predictor coefficients are sent to the LPC composition filter 33.

[0056] LPC composition is performed by the LPC composition filter 33. Thereby, the sound signal of a broadband is compounded. The sound signal compounded with the LPC composition filter 33 is supplied to a band stop filter 36. The output of a band stop filter 36 is supplied to an adder circuit 27.

[0057] In an adder circuit 27, the rise sample circuit 25 and a band pass filter 26 are minded, and the component of the sound signal of the original narrow-band and the component of the sound signal through a band stop filter 36 of a high region from which it synthesized voice are added. Thereby, the sound signal of a broadband is acquired. This sound signal is outputted from an output terminal 37.

[0058] Thus, in the cellular-phone system using VSELP and PSI-CELP as a modulation technique, since linear predictor coefficients and the source of excitation are sent, it can do [extending speech bandwidth or] using such information.

[0059]

[Effect of the Invention] When [which detects a fricative and the affricate] affricate detection is carried out and a fricative and the affricate are detected, he is trying to give a boost to the source of excitation, in case according to this invention LPC composition of the input sound signal is carried out and bandwidth is extended. For this reason, the lack of power when a fricative and the affricate are inputted is improved, and a fricative and the affricate can be reproduced clearly.

[Translation done.]

TECHNICAL FIELD

[Field of the Invention] This invention relates to bandwidth growth equipment and an approach by minding transmission lines, such as the telephone line, from the sound signal with which the frequency band is restricted to the narrow-band at the sound signal processor for generating the sound signal of a broadband and an approach, and a list.

[Translation done.]

PRIOR ART

[Description of the Prior Art] The band of the telephone line is as narrow as 300-3400kHz, and the frequency band of the sound signal sent through the telephone line is restricted. For this reason, the tone quality of the conventional analog telephone line can seldom be said to be fitness. Moreover, there is dissatisfaction also about the tone quality of a digital cellular phone.
[0003] Then, speech bandwidth is extended by the receiver side and the system which aimed at the improvement of tone quality is proposed variously. The narrow-band code book with which the parameter of the narrow-band sound signal beforehand acquired from the pattern of two or more sound signals in this was stored as a code vector, The broadband code book with which the parameter of the wideband voice signal acquired from the pattern of the same sound signal as this was beforehand stored as a code vector is prepared. By a narrow-band code book's analyzing an input signal, and synthesizing voice using a broadband code book based on this analysis result, speech bandwidth is extended and the system which improved tone quality is proposed.
[0004] That is, a frequency band is restricted [as shown in drawing 6 ,] when transmitting a sound signal through a transmission line like the telephone line, and the sound signal from transmission side 101 minds a transmission line 102. For example, even if there is about 7000Hz of frequency bands of the sound signal from transmission side 101 from 300Hz, the frequency band of the sound signal sent to receiver side 103 is restricted to about 3400Hz from 300Hz by minding a transmission line 102.
[0005] Then, as shown in drawing 7 , the narrow-band code book 105 with which the parameter of the narrow-band sound signal beforehand acquired from the pattern of two or more sound signals was stored as a code vector, and the broadband code book 106 with which the parameter of the wideband voice signal with which it was obtained from the pattern of the same sound signal corresponding to the narrow-band code book 105 was beforehand stored as a code vector are prepared.
[0006] In addition, code books 105 and 106 divide the sound signal of the same broadband into the frame of predetermined die length, for example, form the pattern of two or more sound signals, and are created by analyzing spectrum envelopment for every frame. Namely, the sound signal of a broadband is used for code book creation time, and the sound signal of this broadband is divided into it for every predetermined frame. The spectrum envelopment information when analyzing the sound signal of this broadband with a broadband is stored in the broadband code book 106 as a code vector. The spectrum envelopment information when band-limiting the sound signal of a broadband to 300-3400Hz, and analyzing it is stored in the narrow-band code book 105 as a code vector.
[0007] As spectrum envelopment information stored in the narrow-band code book 105 and the broadband code book 106, LPC KEPUTORAMU is used conventionally. LPC KEPUTORAMU is KEPUTORAMU by linear predictor coefficients, is shown by the following formulas, and is made and called for.
[Equation 1]

$$\begin{cases} c_1 = -\alpha_1 \\ c_n = -\alpha_n - \sum_{m=1}^{n-1} \left(1 - \frac{m}{n}\right) \alpha_m c_{n-m} & (1 < n \leqq p) \\ c_n = -\sum_{m=1}^{p} \left(1 - \frac{m}{n}\right) \alpha_m c_{n-m} \end{cases}$$

α ：線形予測係数
p ：線形予測次数

[0008] In drawing 7 , the sound signal of the narrow-band sent to receiver side 103 from transmission side 101 through a transmission line 102 is first sent to the analysis circuit 104. An input sound signal is divided for every predetermined frame, and spectrum envelopment is called for in the analysis circuit 104. The output of the analysis circuit 104 is sent to the narrow-band code book 105. With the narrow-band code book 105, the spectrum envelopment analyzed in the analysis circuit 104 is compared with the spectrum envelopment information stored in the narrow-band code book 105, and matching processing is performed. And the output of the narrow-band code book 105 is sent to the broadband code book 106, and the spectrum envelopment information which matches most in the narrow-band code book 105, and the spectrum envelopment information on a broadband that it corresponds are read from the broadband code book 106.
[0009] This broadband spectrum envelopment information is sent to the synthetic circuit 107. A sound signal is compounded in the synthetic circuit 107 using the spectrum envelopment information on the broadband read from the broadband code book 106. Since this compounded sound signal is compounded using the broadband code book 106, it turns into a sound signal of a broadband.

[Translation done.]

EFFECT OF THE INVENTION

[Effect of the Invention] When [which detects a fricative and the affricate] affricate detection is carried out and a fricative and the affricate are detected, he is trying to give a boost to the source of excitation, in case according to this invention LPC composition of the input sound signal is carried out and bandwidth is extended. For this reason, the lack of power when a fricative and the affricate are inputted is improved, and a fricative and the affricate can be reproduced clearly.

[Translation done.]

TECHNICAL PROBLEM

[Problem(s) to be Solved by the Invention] As mentioned above, in the conventional speech bandwidth escape system, LPC KEPUTORAMU is used as a code book vector. Moreover, the noise and the pulse train are used as a source of excitation at the time of compounding a sound signal. however, the case where the distortion on audibility and a quantization error use the linear scale of a match comparatively in LPC KEPUTORAMU -- a logarithm -- since a scale is used, the small part of energy is thought as important and the error in the large part of energy becomes large. In order to use for such a speech bandwidth escape system, on audibility, it is desirable to suppress distortion by the vowel part as much as possible. Therefore, LPC KEPUTORAMU cannot necessarily be said to be the optimal thing. Moreover, although the thing possible nearest to the LPC remainder of a broadband must be good about the source of excitation, the conventional method using a noise and a pulse train is far from this.
[0011] Then, it is possible [it] to compound a wideband voice signal by LPC composition, using what carried out the rise sample of the LPC remainder as a source of excitation, using an autocorrelation as a code book vector. an autocorrelation -- a logarithm -- since it is not a scale, it is thought that distortion by the vowel part is improved. However, if the sound signal of a broadband is formed by LPC composition, using what carried out the rise sample of the LPC remainder as a source of excitation, using an autocorrelation as a code book vector, especially, a fricative and the affricate will run short and the problem of becoming an unclear sound will arise. It is thought that this originates in the lack of power of the source of excitation as a seed although it is raised to a cause that prediction of spectrum envelopment is not enough, either.
[0012] That is, in the case of a fricative and the affricate, prediction by LPC composition is performed comparatively well, and the power of the remainder becomes small. However, wideband voice of prediction is inadequate and the power of the remainder does not become small. For this reason, in case the band of a fricative and the affricate is extended, remainder power must also be large equally with it. However, since the remainder is predicted and created from the narrow-band remainder, power is not large enough. For this reason, in the case of a fricative and the affricate, the power of the source of excitation runs short.
[0013] Therefore, in case the purpose of this invention extends speech bandwidth, it is to provide with speech bandwidth growth equipment and an approach the sound signal processor which enabled it to reproduce a fricative and the affricate clearly and an approach, and a list.

[Translation done.]

MEANS

[Means for Solving the Problem] After this invention analyzes an input sound signal and performs signal processing to the analyzed sound signal, it is a sound signal processor characterized by to carry out as [have / an affricate detection means detect the fricative and affricate of an input sound signal in the sound signal processor which compounded the sound signal, and a boost means give a boost to the source of excitation when a fricative and the affricate are detected].
[0015] Moreover, an analysis means by which this invention asks for the parameter of an autocorrelation from an input narrow-band sound signal, The source means forming of excitation which asks for the source of excitation from the LPC remainder of an input narrow-band sound signal, The narrow-band code book with which the parameter of the autocorrelation of the narrow-band sound signal beforehand acquired from the pattern of two or more sound signals was stored, The broadband code book with which the parameter of the autocorrelation of the wideband voice signal beforehand acquired from the pattern of two or more sound signals was stored corresponding to the narrow-band code book, An affricate detection means to detect a fricative and the affricate, and a boost means to give a boost to the source of excitation when a fricative and the affricate are detected, A matching means to compare the parameter of the autocorrelation of the sound signal of an input narrow-band with the parameter of the autocorrelation of the input narrow-band sound signal stored in the narrow-band code book, and to search the optimal parameter, Based on the retrieval result in a matching means, the parameter which corresponds out of the parameter of the autocorrelation of the wideband voice signal stored in the broadband code book is read. It is speech bandwidth growth equipment characterized by having a synthetic means to compound an output wideband voice signal based on this read parameter and source of excitation.
[0016] In this invention, an affricate detection means detects a fricative and the affricate using the value of the autocorrelation of an input sound signal, and the value of frame power.
[0017] Thus, if it is made ** which gives a boost to the source of excitation when [which detects a fricative and the affricate] affricate detection is carried out and a fricative and the affricate are detected, the lack of power in the case of a fricative or the affricate is improved, and a fricative and the affricate can be reproduced clearly.
[0018]
[Embodiment of the Invention] Hereafter, the gestalt of implementation of this invention is explained with reference to a drawing. Drawing 1 shows an example of the speech bandwidth escape system by which this invention was applied. In drawing 1 , the narrow-band sound signal whose sampling frequency a frequency band is 8kHz in 300Hz - 3400Hz is supplied to an input terminal 1. This narrow-band sound signal is supplied to the rise sample circuit 3 while it is supplied to the LPC (Linear Predictive Coding) analysis filter 2.
[0019] The rise sample circuit 3 is for carrying out the rise sample of the sampling frequency to 16kHz from 8kHz. The output of the rise sample circuit 3 is supplied to an adder circuit 5 through the band pass filter 4 of a 300Hz - 3400Hz passband. The path leading to this rise sample circuit 3, a band pass filter 4, and an adder circuit 5 is a path for adding the signal of the component of the original frequency band to the sound signal of a high region from which it synthesized voice, as explained later.
[0020] The LPC analysis filter 2 frame-izes the narrow-band sound signal from an input terminal 1, and performs 10th LPC analysis. The 10th autocorrelation is obtained in process of LPC analysis. This autocorrelation is sent to the affricate detector 7 while it is sent to the narrow-band code book 6. Moreover, the LPC remainder searched for with the LPC analysis filter 2 is sent to the rise sample circuit 8.
[0021] The rise sample of the LPC remainder of the voice of a narrow-band is carried out by the rise sample circuit 8. The output of the rise sample circuit 8 is sent to the LPC composition filter 11 through a low pass filter 9 and boost circuit 10 **. The signal which carried out the rise sample of this LPC remainder, and oppressed the high region is used as a source of excitation at the time of compounding a sound signal so that it may explain later. The boost circuit 10 is for boosting the source of excitation, when the affricate and a fricative are detected, and the amount of boosts of the boost circuit 10 is controlled by the output of the affricate detector 7.
[0022] The 10th autocorrelation information on the narrow-band sound signal beforehand acquired from the pattern of two or more sound signals is stored in the narrow-band code book 6 as a code vector. With the narrow-band code book 6, the autocorrelation obtained from the LPC analysis filter 2 is compared with the autocorrelation information stored in the narrow-band code book 6, and matching processing is performed. And the index of the autocorrelation information which matches most is sent to the broadband code book 12.
[0023] Corresponding to the narrow-band code book 6, the 20th autocorrelation information on the wideband voice signal acquired from the sound signal of the same pattern as the time of creating the narrow-band code book 6 is stored in the broadband code book 12 as a code vector. If the autocorrelation information which matches most with the narrow-band code book 6 is judged, this index will be sent to the broadband code book 12, and the autocorrelation information on the broadband corresponding to the autocorrelation information on the narrow-band judged to match most with the broadband code book 12 will be read.
[0024] An autocorrelation is the parameter of a time domain, is the following, and is made and called for.
[Equation 2]

$$r_\tau = \sum_{t=0}^{N-1-\tau} x_t x_{t-\tau} \qquad (\tau \geqq \phi)$$

$$\{x_t\} = \{x_0, x_1, \cdots x_{N-1}\}$$

N ：音声サンプル数

[0025] Using the wideband voice signal whose sampling frequency is 16kHz and which is 0-8000kHz, the broadband code book 12 is the following, and is made and created. That is, this wideband voice signal is divided into 20m frame in every second of advance for die-length 32 m seconds, and the creation time of the broadband code book 12 is asked for the 20th autocorrelation with each frame. A 8-bit code book is created by the GLA (General Lloyd Algorithm) algorithm using this. Let this be the broadband code book 4. Here, the frame number encoded by the i-th code vector of a broadband code book is Ai(ed).
[0026] The narrow-band code book 6 is the same sound signal as having created the broadband code book 12, and a sampling frequency is created using what was restricted to 300Hz - 3400Hz in a frequency band by 8kHz. It is divided into a frame at the time of day as the time of creating the broadband code book 12 when the sound signal restricted to this narrow-band is the same, and the 10th autocorrelation is called for with each frame. And the center of gravity of the narrow-band autocorrelation of the frame belonging to a frame number Ai is searched for, and it is made to make it correspond to the broadband autocorrelation of the broadband code book of a frame number Ai by making the vector into the i-th code vector of the code book of a narrow-band.
[0027] In drawing 1 , the autocorrelation information on the broadband read from the broadband code book 12 is sent to the autocorrelation-linear-predictor-coefficients conversion circuit 13. Conversion to linear predictor coefficients from an autocorrelation is performed by the autocorrelation-linear-predictor-coefficients conversion circuit 13. These linear predictor coefficients are sent to the LPC composition filter 11.
[0028] The rise sample of the LPC remainder from the LPC analysis filter 2 is carried out in the rise sample circuit 8, distortion is generated in the LPC composition filter 11 by return, and the signal which oppressed the high region side through the low pass filter 9 is supplied to it. With the LPC composition filter 11, the rise sample of this LPC remainder is carried out, and LPC composition is performed by the linear predictor coefficients from the autocorrelation-linear-predictor-coefficients conversion circuit section 13, using what oppressed the distorted high region side by return as a source of excitation. Thereby, the sound signal of a 300Hz - 7000Hz broadband is compounded.
[0029] The sound signal compounded with the LPC composition filter 11 is supplied to a band stop filter 14. A band stop filter 14 removes the signal component of the frequency band of an input narrow-band sound signal. By the band stop filter 14, the 300Hz - 3400Hz signal component contained in the sound signal of the original narrow-band is removed out of the sound signal of a broadband with a frequency of 300Hz - 7000Hz compounded with the LPC composition filter 11. The output of this band stop filter 14 is supplied to an adder circuit 5.
[0030] In an adder circuit 5, the component of the sound signal of the narrow-band of origin with a frequency [through the rise sample circuit 3 and a band pass filter 4] of 300Hz - 3400Hz and the component of the sound signal with a frequency [through a band stop filter 14] of 3400Hz - 7000Hz from which it synthesized voice are added. Thereby, the digital sound signal whose sampling frequency a frequency band is 16kHz in 300-7000Hz is acquired. This digital sound signal is outputted from an output terminal 15.
[0031] Thus, in the speech bandwidth growth equipment to which this invention was applied, an input narrow-band sound signal is analyzed using the narrow-band code book 6, and the sound signal of a broadband is compounded using the broadband code book 12. And an autocorrelation is used as information on a code book. rather than LPC KEPUTORAMU uses as spectrum envelopment information and ******** generally uses LPC KEPUTORAMU conventionally as a result of an experiment -- a logarithm -- it is because it turned out that it is more desirable on audibility to use the autocorrelation which is not a scale. this -- LPC KEPUTORAMU -- a logarithm -- since the scale is used -- the small consonant of power -- in a part, it is thought that it is because the error in the large vowel part of power becomes large relatively although an error becomes small.
[0032] And in the speech bandwidth escape system to which this invention was applied, as a source of excitation, the rise sample of the LPC remainder is carried out, distortion is generated by return, and what oppressed the distorted high region side by return is used. If it does in this way, since original audio power and harmonic structure are saved, engine performance sufficient as a source of excitation is obtained.
[0033] Thus, the sound signal of a 300Hz - 7000Hz good broadband is acquired from the LPC composition filter 11 by carrying out the rise sample of the LPC remainder, using as a source of thing excitation which oppressed the distorted high region side by return, using an autocorrelation as information on code books 6 and 12, and compounding a sound signal.
[0034] Thus, the sound signal of the broadband obtained from the LPC composition filter 11 also includes the signal of the frequency component of the original band, and if the output signal of the LPC composition filter 11 is used as it is in order that distortion may attain to the frequency component of the original band by these processings, the distorted effect of the frequency component of the original band will arise.
[0035] Then, by the band stop filter 14, from the output of the LPC composition filter 11, the frequency component of the band of the origin of 300Hz - 3400Hz is removed, and it is carrying out as [add / the component of the sound signal of the origin of 300Hz - 3400Hz taken out through the band pass filter 4, and the component of the 3400Hz - 7000Hz sound signal compounded with the LPC composition filter 11].
[0036] In addition, in distance count of code book creation time, it may be made to perform weighting processing so that the weight of high order data may become small. That is, in the narrow-band code book 6, weight from the 1st order to the 3rd order is set to "1", weight is set to "0" in the degree beyond it, weight from the 1st order to the 6th order is set to "1" in the broadband code book 12, and weight is set to "0" in the degree beyond it. If it does in this way, it not only can perform saving of memory space, but reappearance of a rough spectral envelope will be thought as important as a property of an autocorrelation parameter, and more quality voice will be obtained.
[0037] By the way, if the sound signal of a broadband is formed by LPC composition by making into the source of excitation what carried out the rise sample of the LPC remainder, and oppressed the high region in this way, using an autocorrelation as a code vector, especially, a fricative and the affricate will run short and it will become an unclear sound. It is thought that this originates mainly in the lack of power of the source of excitation although it is raised to a cause that prediction of spectrum envelopment is not enough, either.
[0038] So, in the system to which this invention was applied, the affricate detector 7 which detects a fricative and the affricate, and the boost circuit 10 which boosts all the bands of the source of excitation or some bands when a fricative and the affricate are detected are formed. The

10th autocorrelation called for with the LPC analysis filter 2 is supplied to the affricate detector 7. It is detected whether in the affricate detector 7, a fricative and the affricate were inputted among this 10th autocorrelation using zero-order frame power, the primary autocorrelation, and the secondary autocorrelation. When a fricative and the affricate are detected in the affricate detector 7, all the bands of the source of excitation or some bands are boosted by the boost circuit 10.

[0039] That is, the case of a vowel, and in the case of affricate [a fricative or], as a result of analyzing the autocorrelation of an input sound signal, it turned out that the following differences are in the physical relationship of a zero-order autocorrelation, i.e., frame power, the primary autocorrelation, and the secondary autocorrelation. That is, when R0 and the primary autocorrelation are made into R1 and the secondary autocorrelation R2, as zero-order frame power is shown in drawing 2 , when an input sound signal is a vowel, the zero-order frame power R0, the primary autocorrelation R1, and the secondary autocorrelation R2 are located in a line on an abbreviation straight line. On the other hand, as shown in drawing 3 , in the case of a fricative or the affricate, the physical relationship of the zero-order frame power R0, the primary autocorrelation R1, and the secondary autocorrelation R2 turns into relation which is located in a line convex. If it judges from this whether the physical relationship of R1 and the secondary autocorrelation R2 is located [power / zero-order / frame] in a line convex in R0 and the primary autocorrelation, detection of a fricative or the affricate can be performed.

[0040] When satisfied with the system to which this invention was applied using this of the following conditions, it is judged that they are a fricative and the affricate.

[0041] Conditions (1)

When R0 is more than constant value, R1 is more than constant value and R1/R2 are below constant value, it is judged that they are a fricative and the affricate.

[0042] Conditions (2)

In R'sO being below constant value more than constant value, and R's1 being below constant value and being 1-R1>R1-R2, it judges that they are a fricative and an explosive sound.

[0043] Conditions (3)

In R'sO being below constant value more than constant value, and (R1-dc) /'s (R0-dc's) being below constant value and being 1-R1>R1-R2, it judges that they are a fricative and an explosive sound. In addition, dc is a fixed value for every frame band.

[0044] When it is judged according to conditions (1) or conditions (2) that they are a fricative and the affricate, 10dB of sources of excitation is boosted, for example. Moreover, when it is judged according to conditions (3) that they are a fricative and the affricate, 5dB of sources of excitation is boosted, for example.

[0045] Moreover, if the source of excitation is boosted in an instant when the above conditions are fulfilled, a sound will change suddenly and sense of incongruity will be given. Then, he is made to carry out smoothing of the boost of the source of excitation for every frame, and is trying for change of a boost of the source of excitation not to be noticeable as the source of excitation does not change rapidly.

[0046] It is clear by experiment that the speech bandwidth escape of a good property is performed by the speech bandwidth escape system by which this invention was applied. That is, drawing 4 shows the experimental result when performing the bandwidth escape of a sound signal using the speech bandwidth escape system by which this invention was applied. Drawing 4 A is the spectrum Fig. of the sound signal of the broadband used as the source. The sound signal used as this source shall be band-limited as shown in drawing 4 B, and the speech bandwidth escape system by which this invention was applied shall perform a bandwidth escape. Drawing 4 C is the sound signal acquired by performing the bandwidth escape of this signal. If drawing 4 A is compared with drawing 4 C, the speech bandwidth escape system by which this invention was applied shows that the bandwidth escape of a sound signal was able to be performed in a remarkable precision.

[0047] In addition, this invention can be used for the tone-quality improvement of the telephone line of an analog, and a tone-quality improvement of a digital cellular phone. Especially, in the digital cellular phone, VSELP and PSI-CELP are used as a modulation technique. In VSELP or PSI-CELP, since linear predictor coefficients and the source of excitation are used, such information can be used in the case of the LPC analysis and LPC composition in a speech bandwidth escape system.

[0048] That is, drawing 5 shows the example of application in a digital cellular phone. As shown in drawing 5 , in a digital cellular phone, a parameter equivalent to the source of excitation, linear-predictor-coefficients alpha1 -alpha10, or this is sent. This source of excitation is supplied to an input terminal 21, and linear predictor coefficients are supplied to an input terminal 22. The source of excitation from an input terminal 21 is sent to the rise sample circuit 24 while it is sent to the LPC composition filter 23. The auto correlation coefficient from an input terminal 22 is sent to the LPC composition filter 23.

[0049] With the LPC composition filter 23, based on the source of excitation from an input terminal 21, the linear predictor coefficients from an input terminal 22 are used, and a sound signal is compounded. The sound signal compounded with the LPC composition filter 23 is supplied to the rise sample circuit 25.

[0050] The rise sample circuit 25 is for carrying out the rise sample of the sampling frequency. The output of the rise sample circuit 25 is supplied to an adder circuit 27 through a band pass filter 26. The path leading to this rise sample circuit 25, a band pass filter 26, and an adder circuit 27 is a path for adding to the sound signal which had the signal of the component of the original frequency band compounded.

[0051] Moreover, linear predictor coefficients are sent to the linear-predictor-coefficients-autocorrelation conversion circuit 28 from the LPC composition filter 23. The linear-predictor-coefficients-autocorrelation conversion circuit 28 changes linear predictor coefficients into an autocorrelation. This autocorrelation is sent to the affricate detector 30 while it is sent to the narrow-band code book 29.

[0052] Moreover, the source of excitation from an input terminal 21 is sent to the rise sample circuit 24. The output of the rise sample circuit 24 is sent to the LPC composition filter 33 through a low pass filter 31 and the boost circuit 32. The boost circuit 32 is for boosting the source of excitation, when the affricate and a fricative are detected, and the amount of boosts of the boost circuit 32 is controlled by the output of the affricate detector 30.

[0053] The autocorrelation information on the narrow-band sound signal beforehand acquired from the pattern of two or more sound signals is stored in the narrow-band code book 29 as a code vector. With the narrow-band code book 29, the autocorrelation from the linear-predictor-coefficients-autocorrelation conversion circuit 28 is compared with the autocorrelation information stored in the narrow-band code book 29, and matching processing is performed. And the index of the autocorrelation information which matches most is sent to the broadband code book 34.

[0054] Corresponding to the narrow-band code book 29, the autocorrelation information on the wideband voice signal acquired from the sound signal of the same pattern as the time of creating the narrow-band code book 29 is stored in the broadband code book 34 as a code vector. If the autocorrelation information which matches most with the narrow-band code book 29 is judged, this index will be sent to the broadband code book 34, and the autocorrelation information on the broadband corresponding to the autocorrelation information on the narrow-band judged to match most with the broadband code book 34 will be read.

[0055] The autocorrelation information on the broadband read from the broadband code book 34 is sent to the autocorrelation-linear-predictor-coefficients conversion circuit 35. Conversion to linear predictor coefficients from an autocorrelation is performed by the autocorrelation-linear-predictor-coefficients conversion circuit 35. These linear predictor coefficients are sent to the LPC composition filter 33.
[0056] LPC composition is performed by the LPC composition filter 33. Thereby, the sound signal of a broadband is compounded. The sound signal compounded with the LPC composition filter 33 is supplied to a band stop filter 36. The output of a band stop filter 36 is supplied to an adder circuit 27.
[0057] In an adder circuit 27, the rise sample circuit 25 and a band pass filter 26 are minded, and the component of the sound signal of the original narrow-band and the component of the sound signal through a band stop filter 36 of a high region from which it synthesized voice are added. Thereby, the sound signal of a broadband is acquired. This sound signal is outputted from an output terminal 37.
[0058] Thus, in the cellular-phone system using VSELP and PSI-CELP as a modulation technique, since linear predictor coefficients and the source of excitation are sent, it can do [extending speech bandwidth or] using such information.

[Translation done.]

DESCRIPTION OF DRAWINGS

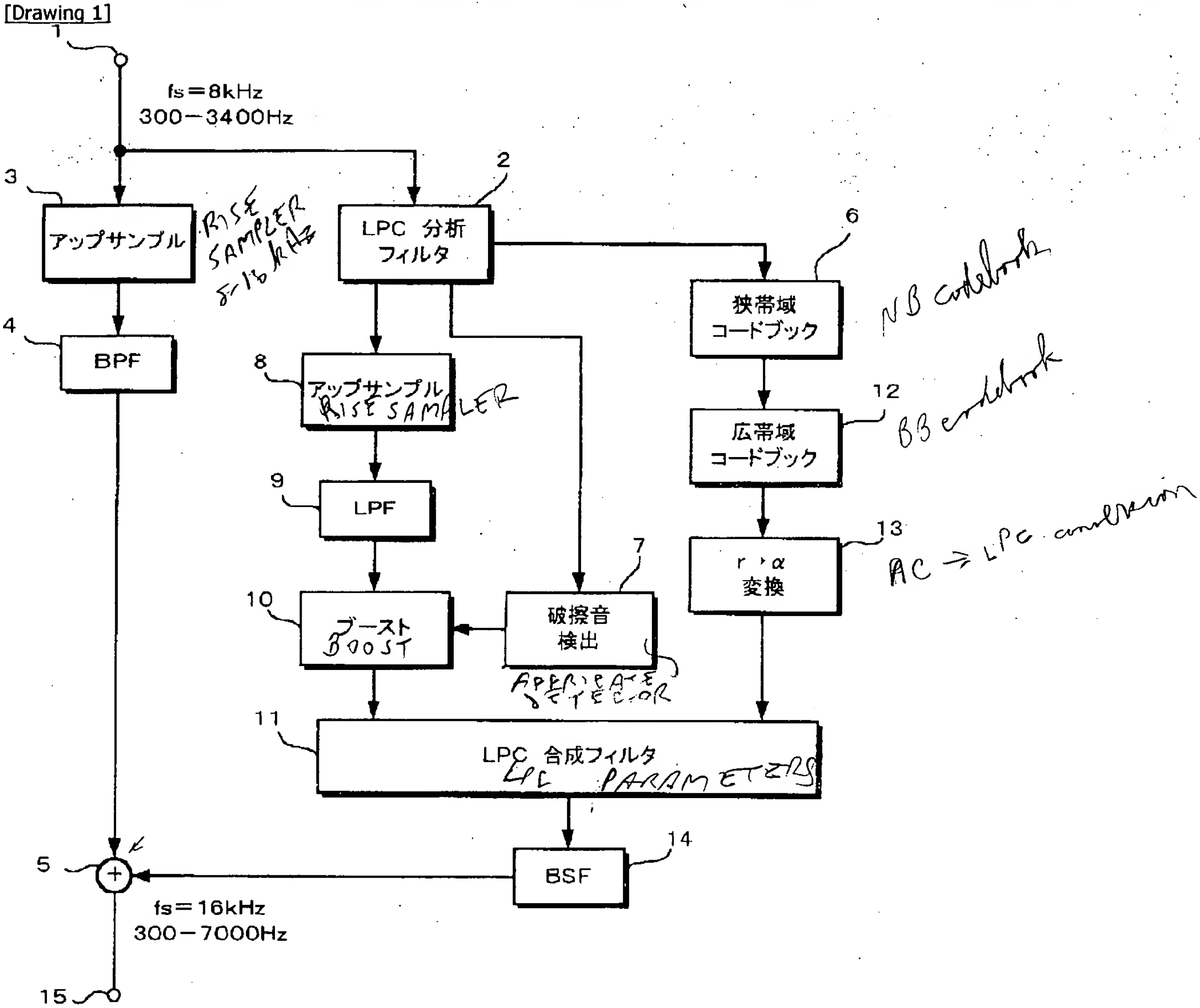
[Brief Description of the Drawings]
[Drawing 1] It is the block diagram showing the speech bandwidth escape structure of a system to which this invention was applied.
[Drawing 2] It is the graph used for explanation of the speech bandwidth escape system by which this invention was applied.
[Drawing 3] It is the graph used for explanation of the speech bandwidth escape system by which this invention was applied.
[Drawing 4] It is the spectrum Fig. used for explanation of the effectiveness of a speech bandwidth escape system that this invention was applied.
[Drawing 5] It is the block diagram showing an example when this invention is applied to a cellular phone.
[Drawing 6] It is the block diagram used for explanation of the voice transmission route to which a frequency band is restricted.
[Drawing 7] It is the block diagram used for explanation of the conventional speech bandwidth escape system.
[Description of Notations]
2 [... An LPC composition filter, 12 / ... Broadband code book] ... An LPC analysis filter, 6 ... A narrow-band code book, 7 ... An affricate detector, 11

[Translation done.]

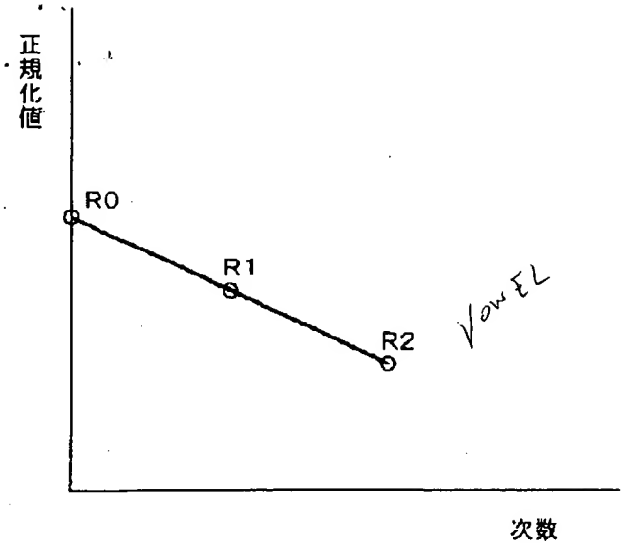
* NOTICES *

DRAWINGS

[Drawing 1]



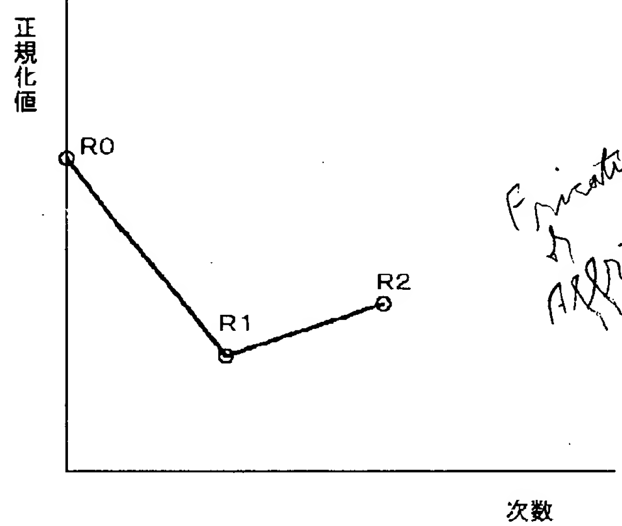
[Drawing 2]



VOWEL

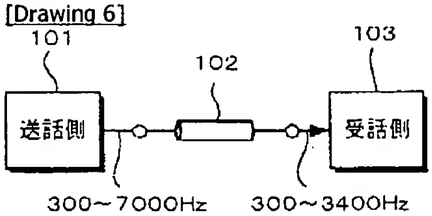
R0
0-order
frame power

[Drawing 3]

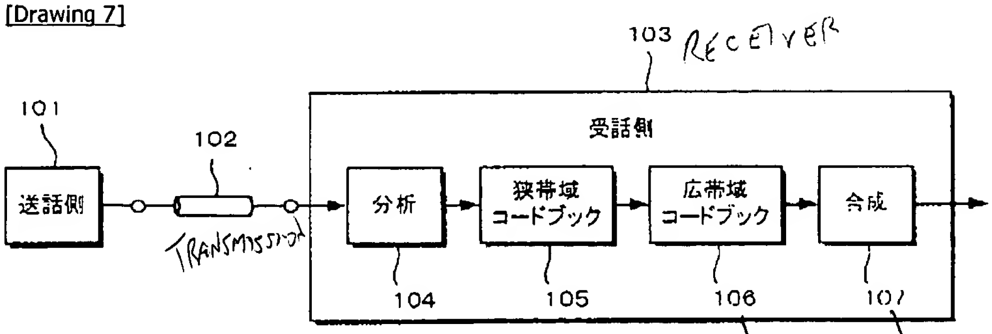


Fricative
&
Affricate

[Drawing 6]



[Drawing 7]



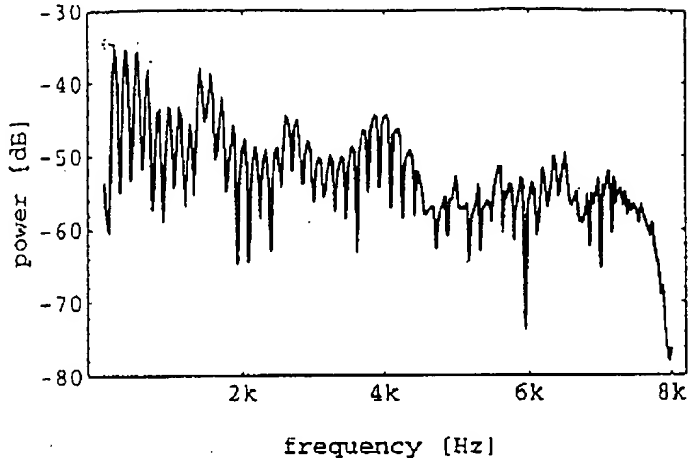
NB CODE BOOK

BB code book

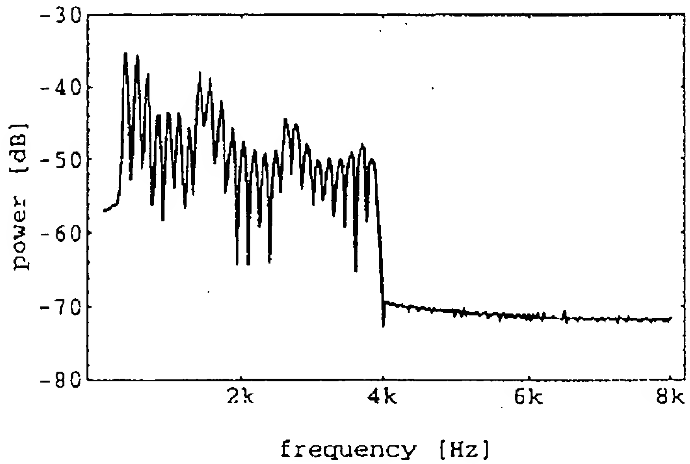
SYNTHESIZER

[Drawing 4]

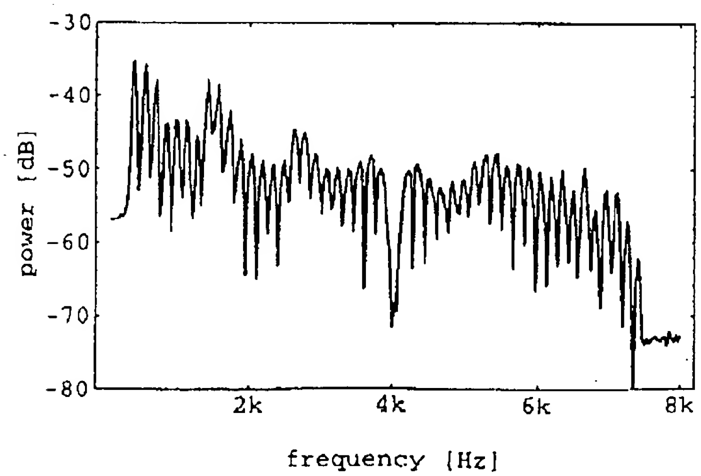
A



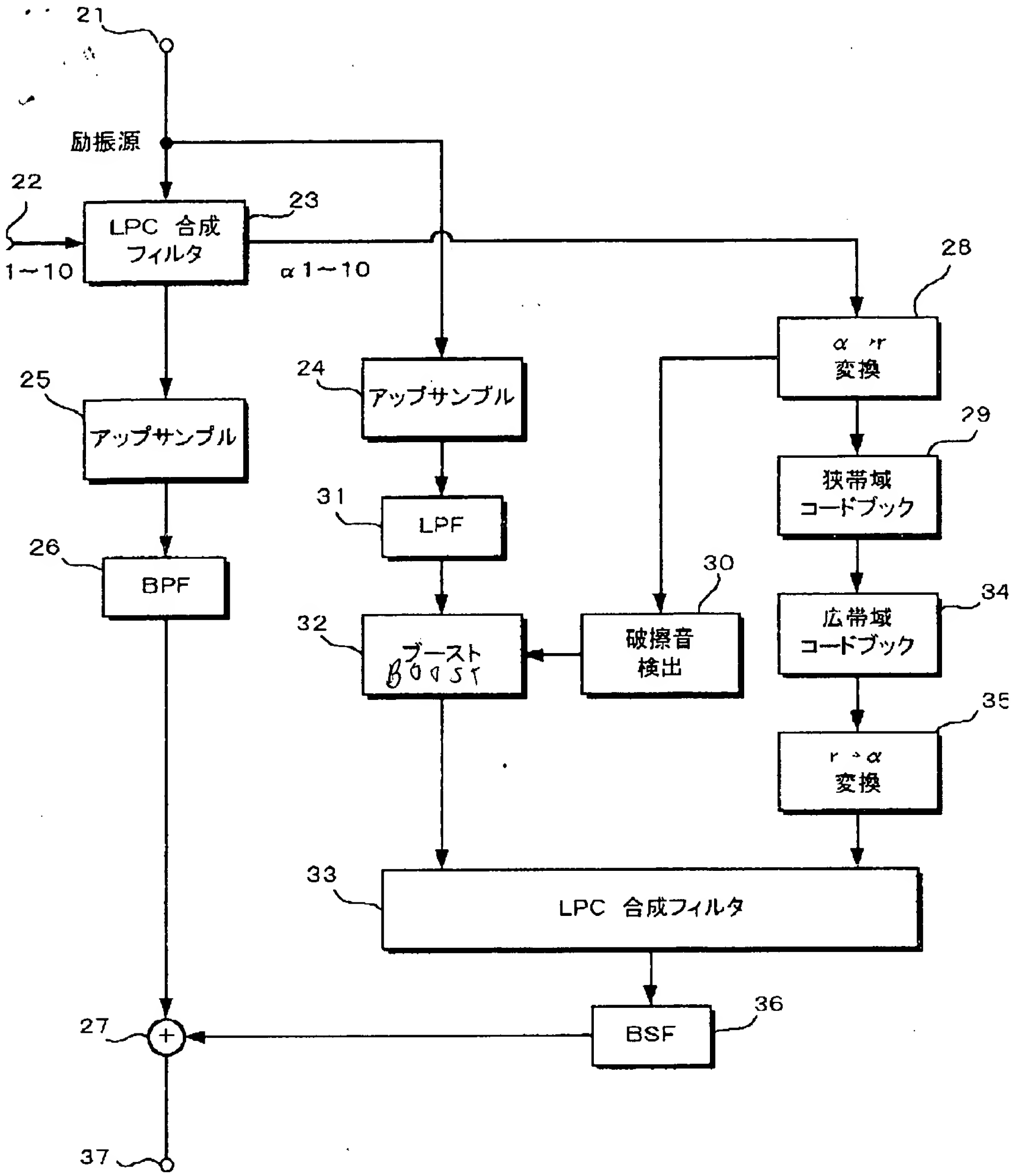
B



C



[Drawing 5]



[Translation done.]

(19)日本国特許庁（ＪＰ）　(12)**公開特許公報**（Ａ）　(11)特許出願公開番号

特開平10－124089

(43)公開日　平成10年(1998)５月15日

(51)Int.Cl.⁶	識別記号	ＦＩ		
Ｇ１０Ｌ　3/02		Ｇ１０Ｌ　3/02		Ｆ
9/08		9/08		Ｃ
Ｈ０４Ｍ　1/00		Ｈ０４Ｍ　1/00		Ｈ

審査請求　未請求　請求項の数12　ＯＬ　（全 11 頁）

(21)出願番号　特願平8－282235

(22)出願日　平成８年(1996)10月24日

(71)出願人　000002185
ソニー株式会社
東京都品川区北品川６丁目７番35号

(72)発明者　大森　士郎
東京都品川区北品川６丁目７番35号　ソニー株式会社内

(72)発明者　西口　正之
東京都品川区北品川６丁目７番35号　ソニー株式会社内

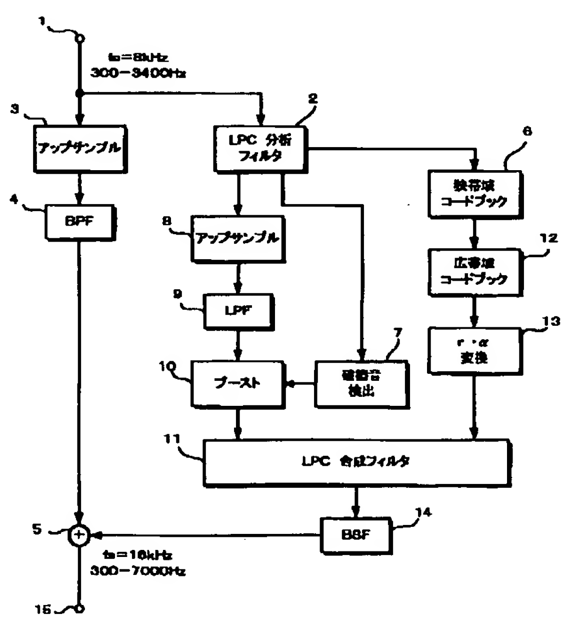
(74)代理人　弁理士　杉浦　正知

(54)【発明の名称】　音声信号処理装置及び方法、並びに、音声帯域幅拡張装置及び方法

(57)【要約】

【課題】　ＬＰＣ合成により音声帯域幅を拡張する際に、摩擦音、破擦音を明瞭に再現できるようにする。

【解決手段】　入力音声信号をＬＰＣ分析フィルタで分析し、自己相関をパラメータとして、狭帯域コードブック６で最もマッチするものを検索し、これに対応するパラメータを広帯域コードブック１２から出力し、ＬＰＣ合成フィルタ１１で合成する。これにより、音声の帯域幅を拡張する。破擦音検出回路７は、入力音声信号の自己相関の値及びフレームパワーの値を用いて、摩擦音や破擦音を検出する。摩擦音や破擦音が検出されたら、励振源の全帯域又は一部の帯域をブーストする。これにより、摩擦音や破擦音の場合のパワー不足が改善され、摩擦音、破擦音を明瞭に再現できる。



1　2

【特許請求の範囲】

【請求項１】　入力音声信号を分析し、上記分析された音声信号に対して信号処理を行なった後、音声信号を合成するようにした音声信号処理装置において、
上記入力音声信号の摩擦音や破擦音を検出する破擦音検出手段と、
上記摩擦音や破擦音が検出された場合に、励振源に対してブーストを与えるブースト手段とを備えるようしたことを特徴とする音声信号処理装置。

【請求項２】　上記破擦音検出手段は、少なくとも上記入力音声信号の自己相関の値及びフレームパワーの値を用いて、摩擦音や破擦音を検出するものである請求項１記載の音声信号処理装置。

【請求項３】　上記ブースト手段は、ブースト値を除々に変化させるものである請求項１又は２記載の音声信号処理装置。

【請求項４】　入力音声信号を分析し、上記分析された音声信号に対して信号処理を行なった後、音声信号を合成するようにした音声信号処理方法において、
上記入力音声信号の摩擦音や破擦音を検出し、
上記摩擦音や破擦音が検出された場合に、励振源の全帯域又は一部の帯域をブーストするようしたことを特徴とする音声信号処理方法。

【請求項５】　上記破擦音の検出は、少なくとも上記入力音声信号の自己相関の値及びフレームパワーの値を用いて行なうようにした請求項４記載の音声信号処理方法。

【請求項６】　上記ブースト値を除々に変化させるようにした請求項４又は５記載の音声信号処理方法。

【請求項７】　入力狭帯域音声信号からパラメータを求める分析手段と、
上記入力狭帯域音声信号のＬＰＣ残差から励振源を求める励振源形成手段と、
予め複数の音声信号のパターンから得られた狭帯域音声信号のパラメータが格納された狭帯域コードブックと、
予め複数の音声信号のパターンから得られた広帯域音声信号のパラメータが上記狭帯域コードブックに対応して格納された広帯域コードブックと、
摩擦音、破擦音を検出する破擦音検出手段と、
上記摩擦音、破擦音が検出されたときに上記励振源に対してブーストを与えるブースト手段と、
上記入力狭帯域の音声信号のパラメータと、上記狭帯域コードブックに格納されている入力狭帯域音声信号のパラメータとを比較し、最適なパラメータを検索するマッチング手段と、
上記マッチング手段での検索結果に基づいて、上記広帯域コードブックに格納されている広帯域音声信号のパラメータの中から対応するパラメータを読み出し、上記読み出されたパラメータと上記励振源を基にして出力広帯域音声信号を合成する合成手段とを備えたことを特徴とする音声帯域幅拡張装置。

【請求項８】　上記破擦音検出手段は、少なくとも上記入力音声信号の自己相関の値及びフレームパワーの値を用いて、摩擦音や破擦音を検出するものである請求項７記載の音声帯域幅拡張装置。

【請求項９】　上記ブースト手段は、ブースト値を除々に変化させるものである請求項７又は８記載の音声帯域幅拡張装置。

【請求項１０】　予め複数の音声信号のパターンから得られた狭帯域音声信号のパラメータが格納された狭帯域コードブックと、予め複数の音声信号のパターンから得られた広帯域音声信号のパラメータが上記狭帯域コードブックに対応して格納された広帯域コードブックとを設け、
入力狭帯域音声信号からパラメータを求める分析し、
上記入力狭帯域音声信号のＬＰＣ残差から励振源を求め、
摩擦音、破擦音を検出し、
上記摩擦音、破擦音が検出されたときに上記励振源に対してブーストを与え、
上記入力狭帯域の音声信号のパラメータと、上記狭帯域コードブックに格納されている入力狭帯域音声信号のパラメータとを比較し、最適なパラメータを検索し、
上記マッチングでの検索結果に基づいて、上記広帯域コードブックに格納されている広帯域音声信号のパラメータの中から対応するパラメータを読み出し、
上記読み出されたパラメータと上記励振源を基にして出力広帯域音声信号を合成するようにしたことを特徴とする音声帯域幅拡張方法。

【請求項１１】　上記破擦音や摩擦音は、上記入力音声信号の自己相関の値及びフレームパワーの値を用いて検出するものである請求項１０記載の音声帯域幅拡張方法。

【請求項１２】　上記ブースト値を除々に変化させるようにした請求項１０又は１１記載の音声帯域幅拡張方法。

【発明の詳細な説明】

【０００１】

【発明の属する技術分野】この発明は、電話回線等の伝送路を介されることにより周波数帯域が狭帯域に制限されている音声信号から広帯域の音声信号を生成するための音声信号処理装置及び方法、並びに、帯域幅拡張装置及び方法に関する。

【０００２】

【従来の技術】電話回線の帯域は例えば３００～３４００ｋＨｚと狭く、電話回線を介して送られてくる音声信号の周波数帯域は制限されている。このため、従来のアナログ電話回線の音質はあまり良好とは言えない。また、ディジタル携帯電話の音質についても不満がある。

【０００３】そこで、受話側で音声帯域幅を拡張し、音

質の改善を図るようにしたシステムが種々提案されている。この中で、予め複数の音声信号のパターンから得られた狭帯域音声信号のパラメータがコードベクタとして格納された狭帯域コードブックと、これと同一の音声信号のパターンから得られた広帯域音声信号のパラメータがコードベクタとして予め格納された広帯域コードブックとを用意しておき、入力信号を狭帯域コードブックで分析し、この分析結果に基づいて広帯域コードブックを用いて音声合成を行なうことにより、音声帯域幅を拡張し、音質を改善するようにしたシステムが提案されている。

【0004】つまり、図6に示すように、電話回線のような伝送路を通じて音声信号を伝送する場合、送話側101からの音声信号は、伝送路102を介されることにより周波数帯域が制限される。例えば、送話側101からの音声信号の周波数帯域が300Hzから7000Hz程度あったとしても、伝送路102を介されることにより、受話側103に送られる音声信号の周波数帯域は、例えば300Hzから3400Hz程度に制限される。

【0005】そこで、図7に示すように、予め複数の音声信号のパターンから得られる狭帯域音声信号のパラメータがコードベクタとして格納された狭帯域コードブッ＊ク105と、狭帯域コードブック105に対応して、同一の音声信号のパターンから得られた広帯域音声信号のパラメータがコードベクタとして予め格納された広帯域コードブック106とが用意される。

【0006】なお、コードブック105及び106は、例えば、同一の広帯域の音声信号を所定の長さのフレームに分割して複数の音声信号のパターンを形成し、各フレーム毎にスペクトラム包絡を分析することにより作成される。すなわち、コードブック作成時には、広帯域の音声信号が用いられ、この広帯域の音声信号が所定のフレーム毎に分割される。広帯域コードブック106には、この広帯域の音声信号を広帯域のまま分析したときのスペクトラム包絡情報がコードベクタとして格納される。狭帯域コードブック105には、広帯域の音声信号を例えば300〜3400Hzに帯域制限して分析したときのスペクトラム包絡情報がコードベクタとして格納される。

【0007】狭帯域コードブック105及び広帯域コードブック106に格納するスペクトラム包絡情報としては、従来、LPCケプトラムが用いられている。LPCケプトラムは、線形予測係数によるケプトラムであり、以下の式で示されるようにして求められる。

【数1】

$$
\begin{cases}
c_1 = -\alpha_1 \\
c_n = -\alpha_n - \sum_{m=1}^{n-1} \left(1 - \frac{m}{n}\right) \alpha_m c_{n-m} & (1 < n \leqq p) \\
c_n = -\sum_{m=1}^{p} \left(1 - \frac{m}{n}\right) \alpha_m c_{n-m}
\end{cases}
$$

α：線形予測係数
p：線形予測次数

【0008】図7において、伝送路102を介して送話側101から受話側103に送られてきた狭帯域の音声信号は、先ず、分析回路104に送られる。分析回路104で、入力音声信号が所定のフレーム毎に分けられ、スペクトラム包絡が求められる。分析回路104の出力が狭帯域コードブック105に送られる。狭帯域コードブック105で、分析回路104で分析されたスペクトラム包絡と、狭帯域コードブック105に格納されているスペクトラム包絡情報とが比較され、マッチング処理が行なわれる。そして、狭帯域コードブック105の出力が広帯域コードブック106に送られ、狭帯域コードブック105において最もマッチしているスペクトラム包絡情報と対応する広帯域のスペクトラム包絡情報が、広帯域コードブック106から読み出される。

【0009】この広帯域スペクトラム包絡情報が合成回路107に送られる。合成回路107で、広帯域コードブック106から読み出された広帯域のスペクトラム包※絡情報を用いて、音声信号が合成される。この合成された音声信号は、広帯域コードブック106を用いて合成されるので、広帯域の音声信号となる。

【0010】

【発明が解決しようとする課題】上述のように、従来の音声帯域幅拡張システムでは、コードブックベクタとしてLPCケプトラムを用いている。また、音声信号を合成する際の励振源としては、ノイズとパルス列を用いている。ところが、LPCケプトラムでは、聴感上の歪と量子化誤差が比較的一致するものの、リニアスケールを用いる場合より対数スケールが用いられるため、エネルギーの小さい部分が重視され、エネルギーの大きい部分での誤差が大きくなる。このような音声帯域幅拡張システムに用いるには、聴感上では、母音部分での歪をできるだけ抑えることが好ましい。したがって、LPCケプトラムは、必ずしも最適なものとは言えない。また、励振源については、広帯域のLPC残差にできるだけ近い

ものが良いはずであるが、ノイズとパルス列を用いた従来の方式は、これとは程遠いものである。

【0011】そこで、コードブックベクタとして自己相関を用い、ＬＰＣ残差をアップサンプルしたものを励振源として用い、ＬＰＣ合成により広帯域音声信号を合成することが考えられる。自己相関は、対数スケールではないので、母音部分での歪が改善されると考えられる。ところが、コードブックベクタとして自己相関を用い、ＬＰＣ残差をアップサンプルしたものを励振源として用い、ＬＰＣ合成により広帯域の音声信号を形成するようにすると、特に、摩擦音、破擦音が不足し、歯切れの悪い音になるという問題が生じる。これは、スペクトラム包絡の予測が十分でないことも原因に上げられるが、種として、励振源のパワー不足に起因すると考えられる。

【0012】すなわち、摩擦音、破擦音の場合、比較的良くＬＰＣ合成による予測が行なわれ、残差のパワーが小さくなる。ところが、広帯域音声では予測が不十分で、残差のパワーが小さくならない。このため、摩擦音、破擦音の帯域を拡張する際には、残差パワーもそれと同等に大きくなっていなければならない。ところが、残差は狭帯域残差から予測して作成されるため、パワーが十分に大きくなっていない。このため、摩擦音、破擦音の場合に、励振源のパワーが不足する。

【0013】したがって、この発明の目的は、音声帯域幅を拡張する際に、摩擦音、破擦音を明瞭に再現できるようにした音声信号処理装置及び方法、並びに、音声帯域幅拡張装置及び方法を提供することにある。

【0014】

【課題を解決するための手段】この発明は、入力音声信号を分析し、分析された音声信号に対して信号処理を行なった後、音声信号を合成するようにした音声信号処理装置において、入力音声信号の摩擦音や破擦音を検出する破擦音検出手段と、摩擦音や破擦音が検出された場合に、励振源に対してブーストを与えるブースト手段とを備えるようしたことを特徴とする音声信号処理装置である。

【0015】また、この発明は、入力狭帯域音声信号から自己相関のパラメータを求める分析手段と、入力狭帯域音声信号のＬＰＣ残差から励振源を求める励振源形成手段と、予め複数の音声信号のパターンから得られた狭帯域音声信号の自己相関のパラメータが格納された狭帯域コードブックと、予め複数の音声信号のパターンから得られた広帯域音声信号の自己相関のパラメータが狭帯域コードブックに対応して格納された広帯域コードブックと、摩擦音、破擦音を検出する破擦音検出手段と、摩擦音、破擦音が検出されたときに励振源に対してブーストを与えるブースト手段と、入力狭帯域の音声信号の自己相関のパラメータと、狭帯域コードブックに格納されている入力狭帯域音声信号の自己相関のパラメータとを比較し、最適なパラメータを検索するマッチング手段と、マッチング手段での検索結果に基づいて、広帯域コードブックに格納されている広帯域音声信号の自己相関のパラメータの中から対応するパラメータを読み出し、この読み出されたパラメータと励振源を基にして出力広帯域音声信号を合成する合成手段とを備えたことを特徴とする音声帯域幅拡張装置である。

【0016】この発明では、破擦音検出手段は、入力音声信号の自己相関の値及びフレームパワーの値を用いて、摩擦音や破擦音を検出するものである。

【0017】このように、摩擦音や破擦音を検出する破擦音検出し、摩擦音や破擦音が検出された場合に、励振源に対してブーストを与えるよにすると、摩擦音や破擦音の場合のパワー不足が改善され、摩擦音、破擦音を明瞭に再現できる。

【0018】

【発明の実施の形態】以下、この発明の実施の形態について図面を参照して説明する。図１は、この発明が適用された音声帯域幅拡張システムの一例を示すものである。図１において、入力端子１に、周波数帯域が例えば３００Ｈｚ～３４００Ｈｚで、サンプリング周波数が８ｋＨｚの狭帯域音声信号が供給される。この狭帯域音声信号は、ＬＰＣ（Linear Predictive Coding）分析フィルタ２に供給されると共に、アップサンプル回路３に供給される。

【0019】アップサンプル回路３は、サンプリング周波数を８ｋＨｚから１６ｋＨｚにアップサンプルするためのものである。アップサンプル回路３の出力は、３００Ｈｚ～３４００Ｈｚの通過帯域のバンドパスフィルタ４を介して、加算回路５に供給される。このアップサンプル回路３、バンドパスフィルタ４、加算回路５に通じる経路は、後に説明するように、元の周波数帯域の成分の信号を、音声合成された高域の音声信号に付加するための経路である。

【0020】LPC分析フィルタ２は、入力端子１からの狭帯域音声信号をフレーム化し、１０次のＬＰＣ分析を行なうものである。ＬＰＣ分析の過程で、１０次の自己相関が得られる。この自己相関は狭帯域コードブック６に送られると共に、破擦音検出回路７に送られる。また、ＬＰＣ分析フィルタ２で求められたＬＰＣ残差は、アップサンプル回路８に送られる。

【0021】アップサンプル回路８により、狭帯域の音声のＬＰＣ残差がアップサンプルされる。アップサンプル回路８の出力がローパスフィルタ９、ブースト回路１０をを介して、ＬＰＣ合成フィルタ１１に送られる。このＬＰＣ残差をアップサンプルし、高域を抑圧した信号は、後に説明するように、音声信号を合成する際の励振源として用いられる。ブースト回路１０は、破擦音や摩擦音が検出された場合に、励振源をブーストするためのもので、ブースト回路１０のブースト量は、破擦音検出回路７の出力により制御される。

【0022】狭帯域コードブック６には、予め複数の音声信号のパターンから得られた狭帯域音声信号の１０次の自己相関情報がコードベクタとして格納されている。狭帯域コードブック６で、ＬＰＣ分析フィルタ２から得られた自己相関と、狭帯域コードブック６に格納されている自己相関情報とが比較され、マッチング処理が行なわれる。そして、最もマッチしている自己相関情報のインデックスが広帯域コードブック１２に送られる。

【0023】広帯域コードブック１２には、狭帯域コードブック６と対応して、狭帯域コードブック６を作成したときと同一のパターンの音声信号から得られる広帯域音声信号の２０次の自己相関情報がコードベクタとして格納されている。狭帯域コードブック６で最もマッチしている自己相関情報が判断されると、このインデックスが広帯域コードブック１２に送られ、広帯域コードブック１２により、最もマッチしていると判断された狭帯域の自己相関情報に対応する広帯域の自己相関情報が読み出される。

【0024】自己相関は、時間領域のパラメータで、以下のようにして求められる。

【数２】

$$r_\tau = \sum_{t=0}^{N-1-\tau} x_t x_{t-\tau} \qquad (\tau \geqq \phi)$$

$$\{x_t\} = \{x_0, x_1, \cdots x_{N-1}\}$$

N : 音声サンプル数

【0025】広帯域コードブック１２は、サンプリング周波数が１６ｋＨｚの、０～８０００ｋＨｚの広帯域音声信号を用いて、以下のようにして作成される。すなわち、広帯域コードブック１２の作成時には、この広帯域音声信号が、長さ３２ｍ秒、前進２０ｍ秒毎のフレームに分割され、各フレームで２０次の自己相関が求められる。これを利用して、ＧＬＡ（General Lloyd Algorithm）アルゴリズムにより、８ビットのコードブックが作成される。これが広帯域コードブック４とされる。ここで、広帯域コードブックのｉ番目のコードベクタにエンコードされたフレーム番号をＡｉする。

【0026】狭帯域コードブック６は、広帯域コードブック１２を作成したのと同一の音声信号で、サンプリング周波数を８ｋＨｚで周波数帯域を３００Ｈｚ～３４００Ｈｚに制限したものを用いて作成される。この狭帯域に制限された音声信号が、広帯域コードブック１２を作成したときと同じ時刻でフレームに分割され、各フレームで１０次の自己相関が求められる。そして、フレーム番号Ａｉに属するフレームの狭帯域自己相関の重心を求め、そのベクターを狭帯域のコードブックのｉ番目のコードベクタとすることで、フレーム番号Ａｉの広帯域コードブックの広帯域自己相関に対応させるようにする。

【0027】図1において、広帯域コードブック１２から読み出された広帯域の自己相関情報は、自己相関－線形予測係数変換回路１３に送られる。自己相関－線形予測係数変換回路１３により、自己相関から線形予測係数への変換が行なわれる。この線形予測係数がＬＰＣ合成フィルタ１１に送られる。

【0028】ＬＰＣ合成フィルタ１１には、ＬＰＣ分析フィルタ２からのＬＰＣ残差をアップサンプル回路８でアップサンプルして折返し歪を発生させ、ローパスフィルタ９を介して高域側を抑圧した信号が供給されている。ＬＰＣ合成フィルタ１１で、このＬＰＣ残差をアップサンプルし、折返し歪の高域側を抑圧したようなものを励振源として用い、自己相関－線形予測係数変換回路部１３からの線形予測係数により、ＬＰＣ合成が行なわれる。これにより、３００Ｈｚ～７０００Ｈｚの広帯域の音声信号が合成される。

【0029】ＬＰＣ合成フィルタ１１で合成された音声信号は、バンドストップフィルタ１４に供給される。バンドストップフィルタ１４は、入力狭帯域音声信号の周波数帯域の信号成分を除去するものである。バンドストップフィルタ１４で、ＬＰＣ合成フィルタ１１で合成された周波数３００Ｈｚ～７０００Ｈｚの広帯域の音声信号の中から、元の狭帯域の音声信号に含まれる３００Ｈｚ～３４００Ｈｚの信号成分が除去される。このバンドストップフィルタ１４の出力が加算回路５に供給される。

【0030】加算回路５で、アップサンプル回路３、バンドパスフィルタ４を介された周波数３００Ｈｚ～３４００Ｈｚの元の狭帯域の音声信号の成分と、バンドストップフィルタ１４を介された周波数３４００Ｈｚ～７０００Ｈｚの音声合成された音声信号の成分とが加算される。これにより、周波数帯域が３００～７０００Ｈｚで、サンプリング周波数が１６ｋＨｚのディジタル音声信号が得られる。このディジタル音声信号が出力端子１５から出力される。

【0031】このように、この発明が適用された音声帯域幅拡張装置では、狭帯域コードブック６を用いて入力狭帯域音声信号が分析され、広帯域コードブック１２を用いて広帯域の音声信号が合成される。そして、コードブックの情報としては、自己相関が用いられる。従来、一般には、スペクトラム包絡情報としてＬＰＣケプトラムが用いらていたが、実験の結果、ＬＰＣケプトラムを用いるより、対数スケールでない自己相関を用いた方が聴感上好ましいことが分かったからである。これは、ＬＰＣケプトラムでは、対数スケールを用いているため、パワーの小さい子音部分では誤差は小さくなるが、パワーの大きい母音部分での誤差が相対的に大きくなるためであると考えられる。

【0032】そして、この発明が適用された音声帯域幅拡張システムでは、励振源として、LPC残差をアップサンプルし、折返し歪を発生させ、折返し歪の高域側を抑圧したものが用いられる。このようにすると、元の音声のパワーや調波構造が保存されているため、励振源として十分な性能が得られる。

【0033】このように、コードブック6、12の情報として自己相関を用い、LPC残差をアップサンプルし、折返し歪の高域側を抑圧したもの励振源として用いて音声信号を合成することにより、LPC合成フィルタ11からは、300Hz～7000Hzの良好な広帯域の音声信号が得られる。

【0034】このようにして、LPC合成フィルタ11から得られる広帯域の音声信号は、元の帯域の周波数成分の信号をも含んでおり、これらの処理により元の帯域の周波数成分に歪が及ぶため、LPC合成フィルタ11の出力信号をそのまま用いると、元の帯域の周波数成分の歪の影響が生じる。

【0035】そこで、バンドストップフィルタ14により、LPC合成フィルタ11の出力から、300Hz～3400Hzの元の帯域の周波数成分を除去し、バンドパスフィルタ4を介して取り出された300Hz～3400Hzの元の音声信号の成分と、LPC合成フィルタ11で合成された3400Hz～7000Hzの音声信号の成分とを加算するようしている。

【0036】なお、コードブック作成時の距離計算において、高次のデータの重みが小さくなるように重み付け処理を行なうようにしても良い。すなわち、狭帯域コードブック6においては1次から3次までの重みを「1」とし、それ以上の次数では重みを「0」とし、広帯域コードブック12においては1次から6次までの重みを「1」とし、それ以上の次数では重みを「0」とする。このようにすると、メモリ容量の節約ができるばかりでなく、自己相関パラメータの性質として、大まかなスペクトル包絡の再現を重視することになり、より品質の良い音声が得られる。

【0037】ところで、このように、コードベクタとして自己相関を用い、LPC残差をアップサンプルして高域を抑圧したものを励振源として、LPC合成により広帯域の音声信号を形成するようにすると、特に、摩擦音、破擦音が不足し、歯切れの悪い音になる。これは、スペクトラム包絡の予測が十分でないことも原因に上げられるが、主として、励振源のパワー不足に起因すると考えられる。

【0038】そこで、この発明が適用されたシステムでは、摩擦音や破擦音を検出する破擦音検出回路7と、摩擦音や破擦音が検出されたときに、励振源の全帯域又は一部の帯域をブーストするブースト回路10が設けられる。破擦音検出回路7には、LPC分析フィルタ2で求められた10次の自己相関が供給される。破擦音検出回路7で、この10次の自己相関のうち、0次のフレームパワー、1次の自己相関、2次の自己相関を用いて、摩擦音や破擦音が入力されたかどうかが検出される。破擦音検出回路7で摩擦音や破擦音が検出されると、ブースト回路10により、励振源の全帯域又は一部の帯域がブーストされる。

【0039】つまり、入力音声信号の自己相関を解析した結果、母音の場合と摩擦音や破擦音の場合とでは、0次の自己相関すなわちフレームパワー、1次の自己相関、2次の自己相関の位置関係に、以下のような違いがあることが分かった。すなわち、0次のフレームパワーをR0、1次の自己相関をR1、2次の自己相関R2とすると、図2に示すように、入力音声信号が母音の場合には、0次のフレームパワーR0、1次の自己相関R1、2次の自己相関R2が略直線上に並ぶ。これに対して、図3に示すように、摩擦音や破擦音の場合には、0次のフレームパワーR0、1次の自己相関R1、2次の自己相関R2の位置関係が、下に凸に並ぶような関係となる。このことから、0次のフレームパワーをR0、1次の自己相関をR1、2次の自己相関R2の位置関係が下に凸に並んでいるかどうかを判断すれば、摩擦音や破擦音の検出が行なえる。

【0040】このことを利用して、この発明が適用されたシステムでは、以下の条件を満足するときに摩擦音や破擦音であると判断される。

【0041】条件（1）

R0が一定値以上であり、かつR1が一定値以上であり、かつR1／R2が一定値以下である場合には、摩擦音や破擦音であると判断する。

【0042】条件（2）

R0が一定値以上、一定値以下であり、かつR1が一定値以下であり、かつ1－R1＞R1－R2である場合には、摩擦音や破裂音であると判断する。

【0043】条件（3）

R0が一定値以上、一定値以下であり、かつ（R1－dc）／（R0－dc）が一定値以下であり、かつ1－R1＞R1－R2である場合には、摩擦音や破裂音であると判断する。なお、dcはフレームバンド毎に一定の値である。

【0044】条件（1）又は条件（2）により摩擦音や破擦音であると判断された場合には、励振源が例えば10dBブーストされる。また、条件（3）により摩擦音や破擦音であると判断された場合には、励振源が例えば5dBブーストされる。

【0045】また、以上のような条件が満たされるときに、瞬時に励振源のブーストを行なってしまうと、急に音が変化して、違和感を与える。そこで、励振源が急激に変化しないように、フレーム毎に励振源のブーストをスムージングするようにし、励振源のブーストの変化が目立たないようにしている。

【0046】この発明が適用された音声帯域幅拡張システムにより、良好な特性の音声帯域幅拡張が行なわれることは、実験により明らかである。すなわち、図4は、この発明が適用された音声帯域幅拡張システムを用いて音声信号の帯域幅拡張を行なったときの実験結果を示すものである。図４Ａは、ソースとなる広帯域の音声信号のスペクトラム図である。このソースとなる音声信号を、図４Ｂに示すように帯域制限し、この発明が適用された音声帯域幅拡張システムにより帯域幅拡張を行なうものとする。図４Ｃは、この信号の帯域幅拡張を行なって得られた音声信号である。図４Ａと図４Ｃとを比較すれば、この発明が適用された音声帯域幅拡張システムにより、かなりの精度で音声信号の帯域幅拡張が行なえたことが分かる。

【0047】なお、この発明は、アナログの電話回線の音質改善や、ディジタル携帯電話の音質改善に用いることができる。特に、ディジタル携帯電話では、変調方式としてＶＳＥＬＰやＰＳＩ－ＣＥＬＰが用いられている。ＶＳＥＬＰやＰＳＩ－ＣＥＬＰでは、線形予測係数や励振源が使われるので、これらの情報を音声帯域幅拡張システムにおけるＬＰＣ分析やＬＰＣ合成の際に用いることができる。

【0048】つまり、図5はディジタル携帯電話での適用例を示すものである。図5に示すように、ディジタル携帯電話においては、励振源と線形予測係数α_1 ～α_{10}若しくはこれと等価なパラメータが送られてくる。この励振源が入力端子２１に供給され、線形予測係数が入力端子２２に供給される。入力端子２１からの励振源は、ＬＰＣ合成フィルタ２３に送られると共に、アップサンプル回路２４に送られる。入力端子２２からの自己相関係数は、ＬＰＣ合成フィルタ２３に送られる。

【0049】ＬＰＣ合成フィルタ２３で、入力端子２１からの励振源を基に、入力端子２２からの線形予測係数を用いて、音声信号が合成される。ＬＰＣ合成フィルタ２３で合成された音声信号は、アップサンプル回路２５に供給される。

【0050】アップサンプル回路２５は、サンプリング周波数をアップサンプルするためのものである。アップサンプル回路２５の出力は、バンドパスフィルタ２６を介して、加算回路２７に供給される。このアップサンプル回路２５、バンドパスフィルタ２６、加算回路２７に通じる経路は、元の周波数帯域の成分の信号を合成された音声信号に付加するための経路である。

【0051】また、ＬＰＣ合成フィルタ２３から線形予測係数－自己相関変換回路２８に線形予測係数が送られる。線形予測係数－自己相関変換回路２８は、線形予測係数を自己相関に変換するものである。この自己相関は狭帯域コードブック２９に送られると共に、破擦音検出回路３０に送られる。

【0052】また、入力端子２１からの励振源は、アップサンプル回路２４に送られる。アップサンプル回路２４の出力がローパスフィルタ３１、ブースト回路３２を介して、ＬＰＣ合成フィルタ３３に送られる。ブースト回路３２は、破擦音や摩擦音が検出された場合に励振源をブーストするためのもので、ブースト回路３２のブースト量は、破擦音検出回路３０の出力により制御される。

【0053】狭帯域コードブック２９には、予め複数の音声信号のパターンから得られた狭帯域音声信号の自己相関情報がコードベクタとして格納されている。狭帯域コードブック２９で、線形予測係数－自己相関変換回路２８からの自己相関と、狭帯域コードブック２９に格納されている自己相関情報とが比較され、マッチング処理が行なわれる。そして、最もマッチしている自己相関情報のインデックスが広帯域コードブック３４に送られる。

【0054】広帯域コードブック３４には、狭帯域コードブック２９と対応して、狭帯域コードブック２９を作成したときと同一のパターンの音声信号から得られる広帯域音声信号の自己相関情報がコードベクタとして格納されている。狭帯域コードブック２９で最もマッチしている自己相関情報が判断されると、このインデックスが広帯域コードブック３４に送られ、広帯域コードブック３４により、最もマッチしていると判断された狭帯域の自己相関情報に対応する広帯域の自己相関情報が読み出される。

【0055】広帯域コードブック３４から読み出された広帯域の自己相関情報は、自己相関－線形予測係数変換回路３５に送られる。自己相関－線形予測係数変換回路３５により、自己相関から線形予測係数への変換が行なわれる。この線形予測係数がＬＰＣ合成フィルタ３３に送られる。

【0056】ＬＰＣ合成フィルタ３３で、ＬＰＣ合成が行なわれる。これにより、広帯域の音声信号が合成される。ＬＰＣ合成フィルタ３３で合成された音声信号は、バンドストップフィルタ３６に供給される。バンドストップフィルタ３６の出力が加算回路２７に供給される。

【0057】加算回路２７で、アップサンプル回路２５、バンドパスフィルタ２６を介され元の狭帯域の音声信号の成分と、バンドストップフィルタ３６を介された音声合成された高域の音声信号の成分とが加算される。これにより、広帯域の音声信号が得られる。この音声信号が出力端子３７から出力される。

【0058】このように、変調方式としてＶＳＥＬＰやＰＳＩ－ＣＥＬＰを用いた携帯電話システムでは、線形予測係数や励振源が送られてくるので、これらの情報を用いて、音声帯域幅を拡張することかできる。

【0059】

【発明の効果】この発明によれば、入力音声信号をＬＰＣ合成して帯域幅を拡張する際に、摩擦音や破擦音を検

出する破擦音検出し、摩擦音や破擦音が検出された場合に、励振源に対してブーストを与えるようにしている。このため、摩擦音や破擦音が入力された場合のパワー不足が改善され、摩擦音、破擦音を明瞭に再現できる。

【図面の簡単な説明】

【図1】この発明が適用された音声帯域幅拡張システムの構成を示すブロック図である。

【図2】この発明が適用された音声帯域幅拡張システムの説明に用いるグラフである。

【図3】この発明が適用された音声帯域幅拡張システムの説明に用いるグラフである。

【図4】この発明が適用された音声帯域幅拡張システムの効果の説明に用いるスペクトラム図である。

【図5】この発明が携帯電話に適用された場合の一例を示すブロック図である。

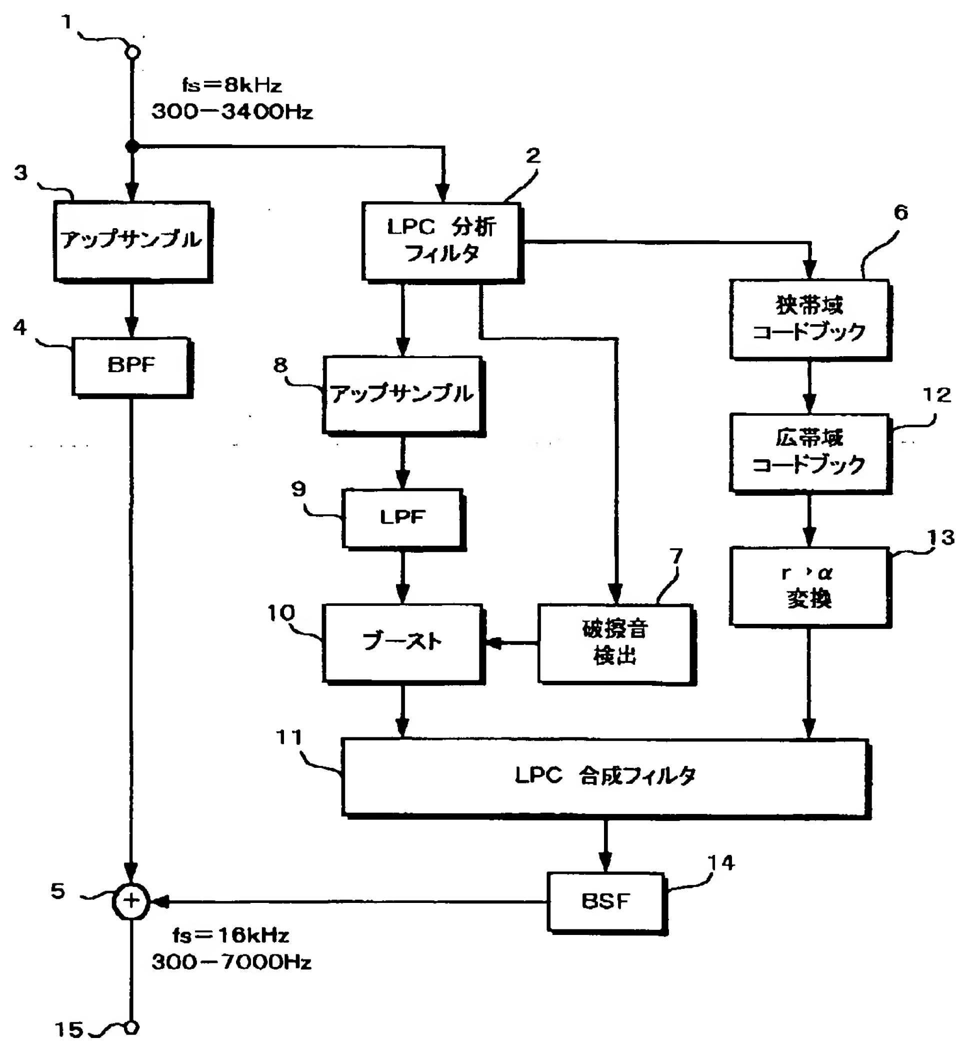
【図6】周波数帯域が制限される音声伝送経路の説明に用いるブロック図である。

【図7】従来の音声帯域幅拡張システムの説明に用いるブロック図である。

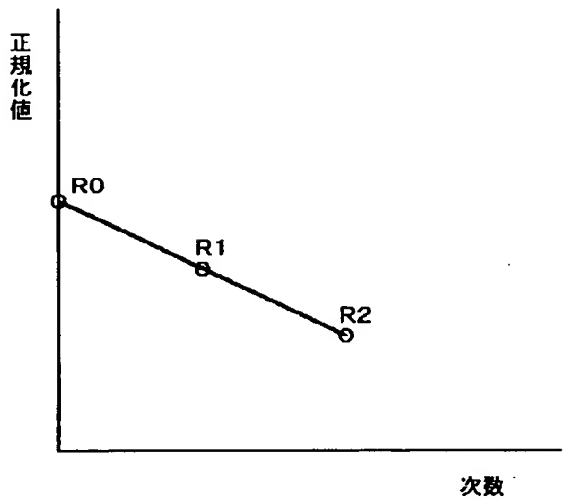
【符号の説明】

2・・・LPC分析フィルタ，6・・・狭帯域コードブック，7・・・破擦音検出回路，11・・・LPC合成フィルタ，12・・・広帯域コードブック

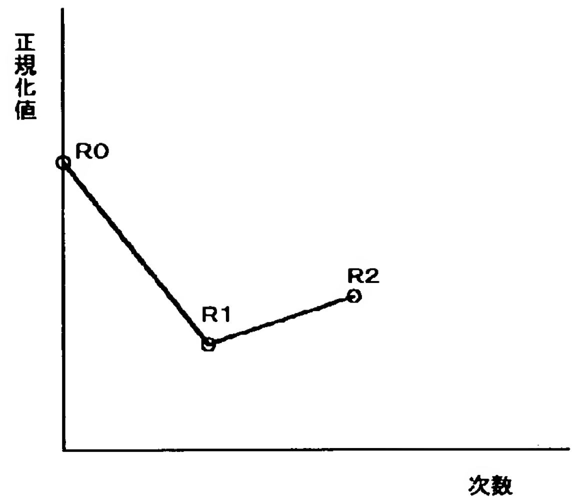
【図1】



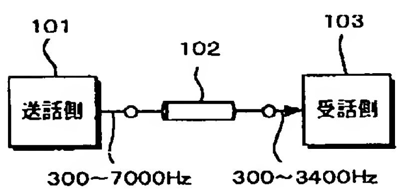
【図2】


正規化値
R0
R1
R2
次数

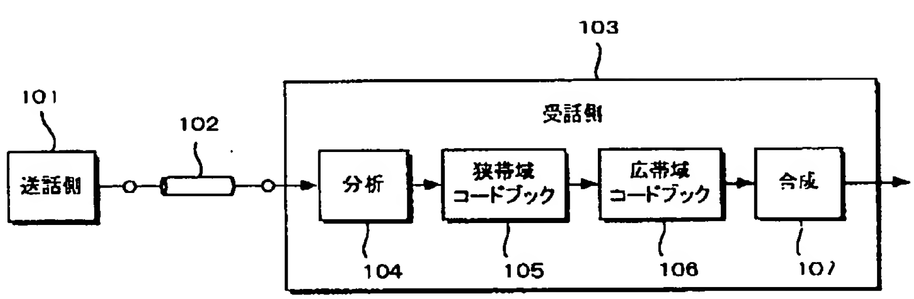
【図3】


正規化値
R0
R1
R2
次数

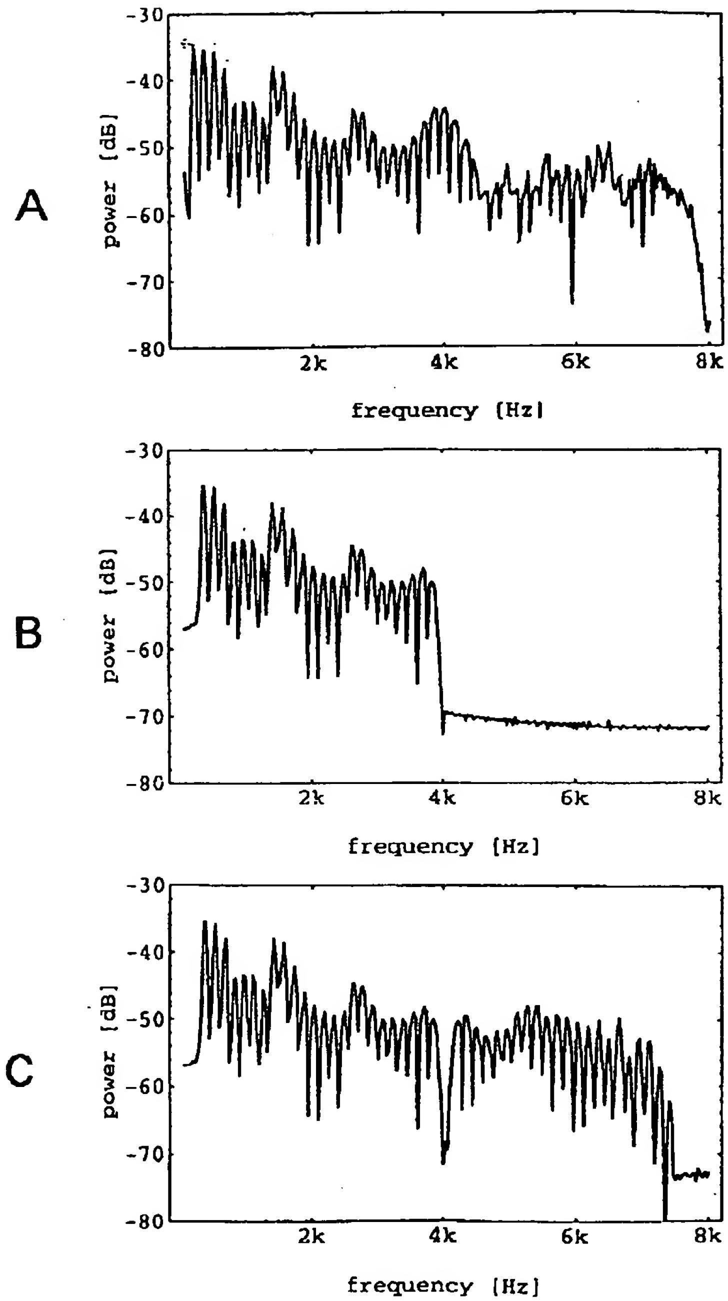
【図6】


101
102
103
送話側
受話側
300～7000Hz
300～3400Hz

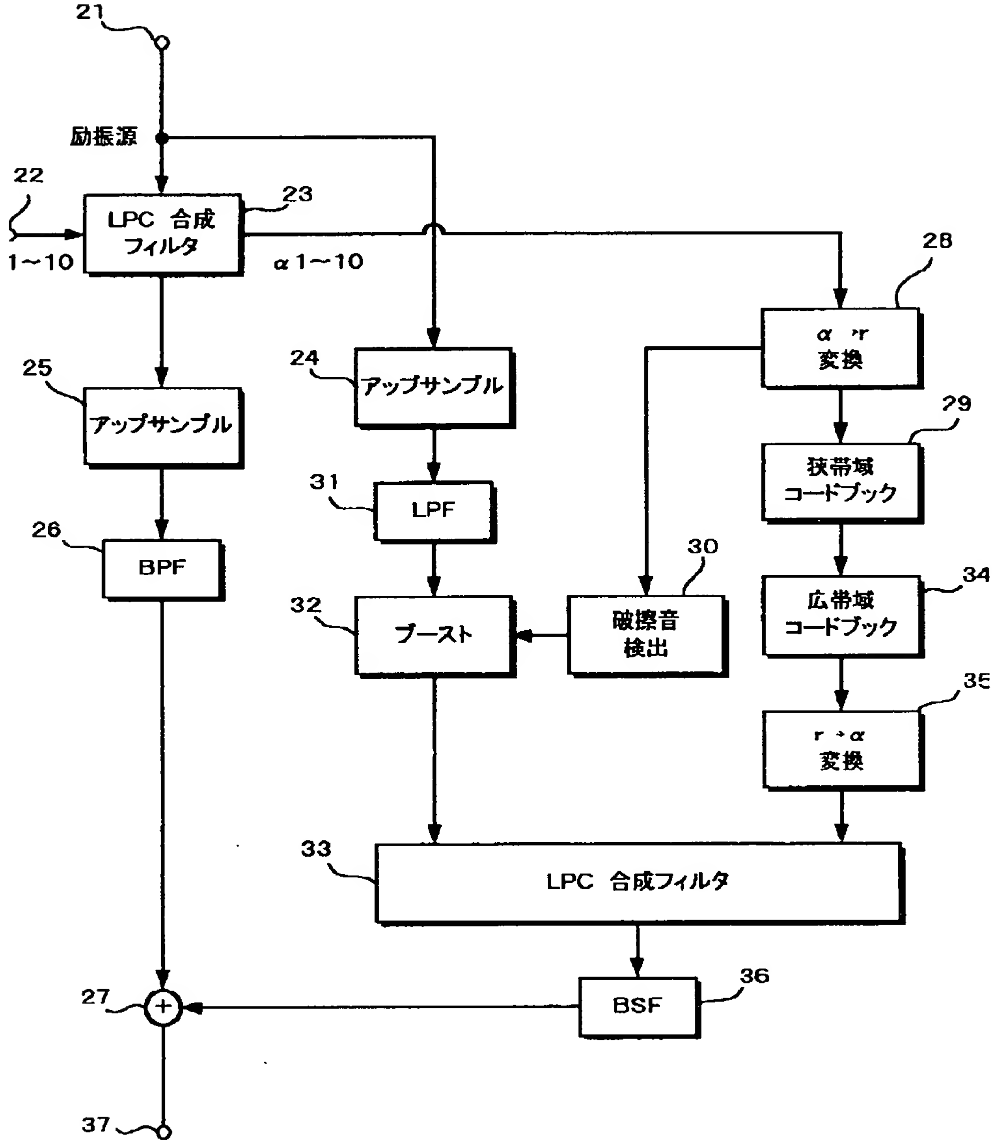
【図7】


101
102
103
受話側
送話側
分析
狭帯域
コードブック
広帯域
コードブック
合成
104
105
106
107

【図4】



【図5】


21
励振源
22
1～10
LPC 合成
フィルタ
23
α1～10
28
α→r
変換
24
アップサンプル
25
アップサンプル
29
狭帯域
コードブック
31
LPF
26
BPF
30
破擦音
検出
34
広帯域
コードブック
32
ブースト
35
r→α
変換
33
LPC 合成フィルタ
36
BSF
27
37